Dokumentvorlage, Version vom 18.04.2013

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Lenvatinib (Lenvima®)

Eisai GmbH

Modul 4B

Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem hepatozellulärem Karzinom (HCC), die zuvor noch keine systemische Therapie erhalten haben

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Stand: 18.09.2018

Inhaltsverzeichnis

Seite

Tabellenverzeichnis ... 4
Eigene Tabellen ... 9
Abbildungsverzeichnis ... 10
Abkürzungsverzeichnis ... 22
4 Modul 4 – allgemeine Informationen ... 26
4.1 Zusammenfassung der Inhalte von Modul 4 ... 27
4.2 Methodik ... 46
4.2.1 Fragestellung ... 46
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung ... 49
4.2.3 Informationsbeschaffung ... 53
4.2.3.1 Studien des pharmazeutischen Unternehmers ... 53
4.2.3.2 Bibliografische Literaturrecherche ... 53
4.2.3.3 Suche in Studienregistern ... 54
4.2.3.4 Selektion relevanter Studien ... 55
4.2.4 Bewertung der Aussagekraft der Nachweise ... 56
4.2.5 Informationssynthese und -analyse ... 59
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien ... 59
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien ... 60
4.2.5.3 Meta-Analysen ... 74
4.2.5.4 Sensitivitätsanalysen ... 75
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren ... 77
4.2.5.6 Indirekte Vergleiche ... 80
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen ... 82
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel ... 82
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel ... 82
4.3.1.1.1 Studien des pharmazeutischen Unternehmers ... 82
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche ... 83
4.3.1.1.3 Studien aus der Suche in Studienregistern ... 85
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel ... 86
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ... 88
4.3.1.2.1 Studiendesign und Studienpopulationen ... 88
4.3.1.2.2 Verzerrungspotenzial auf Studienebene ... 103
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien ... 104
4.3.1.3.1 Endpunkt – RCT ... 104
4.3.1.3.1.1 Endpunkt Mortalität – RCT ... 105
4.3.1.3.1.2 Endpunkt Morbidität – RCT ... 112
4.3.1.3.1.3 Endpunkt Gesundheitsbezogene Lebensqualität – RCT ... 154
4.3.1.3.1.4 Endpunkte zu Sicherheit und Verträglichkeit - RCT ... 175

4.3.1.3.2 Subgruppenanalysen – RCT ... 191
4.3.1.3.2.1 Subgruppenanalyse für den Endpunkt Mortalität ... 193
4.3.1.3.2.2 Subgruppenanalyse für den Endpunkt Morbidität ... 194
4.3.1.3.2.3 Subgruppenanalyse für den Endpunkt Gesundheitsbezogene Lebensqualität ... 223
4.3.1.3.2.4 Subgruppenanalyse für die Endpunkte zu Sicherheit und Verträglichkeit ... 251
4.3.2 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien ... 260
4.3.3 Weitere Unterlagen ... 265
4.3.3.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien ... 265
4.3.3.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche ... 265
4.3.3.1.2 Charakteristika der Studien für indirekte Vergleiche ... 265
4.3.3.1.3 Ergebnisse aus indirekten Vergleichen ... 266
4.3.3.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT ... 266
4.3.3.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT ... 268
4.3.3.2 Nicht randomisierte vergleichende Studien ... 269
4.3.3.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien ... 269
4.3.3.2.2 Charakteristika der nicht randomisierten vergleichenden Studien ... 269
4.3.3.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien ... 270
4.3.3.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 270
4.3.3.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien ... 271
4.3.3.3 Weitere Untersuchungen ... 271
4.3.3.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen ... 271
4.3.3.3.2 harakteristika der weiteren Untersuchungen ... 272
4.3.3.3.3 Ergebnisse aus weiteren Untersuchungen ... 272
4.3.3.3.3.1 <Endpunkt xxx> – weitere Untersuchungen ... 272
4.3.3.3.3.2 Subgruppenanalysen – weitere Untersuchungen ... 273
4.3.3.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen ... 273
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens ... 273
4.4.1 Beurteilung der Aussagekraft der Nachweise ... 273
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß ... 276
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht ... 284
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte ... 284
4.5.1 Begründung für die Vorlage indirekter Vergleiche ... 284
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen ... 285
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen ... 285
4.5.4 Verwendung von Surrogatendpunkten ... 285
4.6 Liste der eingeschlossenen Studien ... 286
4.7 Referenzliste ... 287
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche ... 293
Anhang 4-B : Suchstrategien – Suche in Studienregistern ... 299

Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)..... 301
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)..... 303
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT 314
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 332

Tabellenverzeichnis

Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 83
Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel 83
Tabelle 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel 86
Tabelle 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel 87
Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – Studie 304 mit dem zu bewertenden Arzneimittel 89
Tabelle 4-6: Charakterisierung der Interventionen – Studie 304 mit dem zu bewertenden Arzneimittel 90
Tabelle 4-7: Charakterisierung der Studienpopulationen: demographische Charakteristika zu Studienbeginn – Studie 304 mit dem zu bewertenden Arzneimittel 91
Tabelle 4-8: Charakterisierung der Studienpopulationen: Krankheitsgeschichte zu Studienbeginn – Studie 304 mit dem zu bewertenden Arzneimittel 94
Tabelle 4-9: Vorangegangene Tumorbehandlungen der Patienten im BCLC- Stadium B zu Baseline – Studie 304 mit dem zu bewertenden Arzneimittel 102
Tabelle 4-10: Verzerrungspotenzial auf Studienebene – Studie 304 mit dem zu bewertenden Arzneimittel 103
Tabelle 4-11: Matrix der Endpunkte in der eingeschlossenen Studie 304 mit dem zu bewertenden Arzneimittel 104
Tabelle 4-12: Operationalisierung des Endpunkts OS 106
Tabelle 4-13: Bewertung des Verzerrungspotenzials für den Endpunkt OS in Studie 304 mit dem zu bewertenden Arzneimittel 107
Tabelle 4-14: Hauptanalyse: Ergebnisse für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel 108
Tabelle 4-15: Sensitivitätsanalyse: Ergebnisse der adjustierten Analysen für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel 109
Tabelle 4-16: Postprogressionstherapie: Ergebnisse für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel 111
Tabelle 4-17: Operationalisierung des Endpunkts PFS 113
Tabelle 4-18: Bewertung des Verzerrungspotenzials für den Endpunkt PFS in Studie 304 mit dem zu bewertenden Arzneimittel 114
Tabelle 4-19: Hauptanalyse: Ergebnisse für den Endpunkt PFS (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 115
Tabelle 4-20: Sensitivitätsanalyse: Ergebnisse für den Endpunkt PFS (IIR, RECIST v.1.1) aus Studie 304 mit dem zu bewertenden Arzneimittel 117
Tabelle 4-21: Sensitivitätsanalyse: Ergebnisse für den Endpunkt PFS (Prüfarzt, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 118

Tabelle 4-22: Ergänzende Analyse: Ergebnisse für den Endpunkt ORR (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 120
Tabelle 4-23: Sensitivitätsanalyse: Ergebnisse für den Endpunkt ORR (IIR, RECIST v.1.1) aus Studie 304 mit dem zu bewertenden Arzneimittel 120
Tabelle 4-24: Sensitivitätsanalyse: Ergebnisse für den Endpunkt ORR (Prüfarzt, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 121
Tabelle 4-25: Ergebnisse für die Endpunkte CR und PR aus Studie 304 mit dem zu bewertenden Arzneimittel 121
Tabelle 4-26: Operationalisierung des Endpunkts EORTC QLQ-C30 (Symptome) 124
Tabelle 4-27: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-C30 (Symptome) in Studie 304 mit dem zu bewertenden Arzneimittel 125
Tabelle 4-28: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 126
Tabelle 4-29: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 126
Tabelle 4-30: Hauptanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 131
Tabelle 4-31: Operationalisierung des Endpunkts EORTC QLQ-HCC18 (Symptome) 138
Tabelle 4-32: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-HCC18 (Symptome) in Studie 304 mit dem zu bewertenden Arzneimittel 139
Tabelle 4-33: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 140
Tabelle 4-34: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 140
Tabelle 4-35: Hauptanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 144
Tabelle 4-36: Operationalisierung des Endpunkts EQ-5D VAS 148
Tabelle 4-37: Bewertung des Verzerrungspotenzials für den Endpunkt EQ-5D VAS in Studie 304 mit dem zu bewertenden Arzneimittel 149
Tabelle 4-38: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des Fragebogens EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 150
Tabelle 4-39: Ergebnisse der Baseline-Werte auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 150
Tabelle 4-40: Hauptanalyse: Ergebnisse für Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 152

Tabelle 4-41: Operationalisierung des Endpunkts EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) ... 154
Tabelle 4-42: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) in Studie 304 mit dem zu bewertenden Arzneimittel ... 155
Tabelle 4-43: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des Fragebogens EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 156
Tabelle 4-44: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 156
Tabelle 4-45: Hauptanalyse: Ergebnisse für Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 161
Tabelle 4-46: Operationalisierung des Endpunkts EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) ... 167
Tabelle 4-47: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) in Studie 304 mit dem zu bewertenden Arzneimittel ... 168
Tabelle 4-48: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des Fragebogens EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 169
Tabelle 4-49: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 169
Tabelle 4-50: Hauptanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 172
Tabelle 4-51: Operationalisierung der Endpunkte zu Sicherheit und Verträglichkeit ... 175
Tabelle 4-52: Bewertung des Verzerrungspotenzials für die Endpunkte zu Sicherheit und Verträglichkeit[a] in Studie 304 mit dem zu bewertenden Arzneimittel ... 177
Tabelle 4-53: Zusammenfassung der Länge der Gesamtbeobachtungszeit der UE (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 178
Tabelle 4-54: Hauptanalyse: Ergebnisse für die Zeit bis zum Auftreten des ersten UE (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 181
Tabelle 4-55: Ergänzende Analyse: Ergebnisse für die Zeit bis zum Auftreten des ersten UE mit CTCAE Grad 3 oder 4 (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel ... 189
Tabelle 4-56: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel ... 193

Tabelle 4-57: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für den Endpunkt PFS (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 194

Tabelle 4-58: Subgruppenanalyse: Ergebnisse für den Endpunkt PFS (IIR, mRECIST) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 196

Tabelle 4-59: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 197

Tabelle 4-60: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 210

Tabelle 4-61: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel 215

Tabelle 4-62: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 218

Tabelle 4-63: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 220

Tabelle 4-64: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 222

Tabelle 4-65: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel 223

Tabelle 4-66: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 238

Tabelle 4-67: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel 242

Tabelle 4-68: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 248

Tabelle 4-69: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zum Auftreten des ersten UE (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel 251
Tabelle 4-70: Subgruppenanalyse: Ergebnisse für die Zeit bis zum Auftreten des ersten UE (in Tagen) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel 258
Tabelle 4-71: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche 266
Tabelle 4-72: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 266
Tabelle 4-73: Operationalisierung von <Endpunkt xxx> 267
Tabelle 4-74: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche 267
Tabelle 4-75: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche 268
Tabelle 4-76: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien 270
Tabelle 4-77: Operationalisierung von <Endpunkt xxx> 270
Tabelle 4-78: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien 271
Tabelle 4-79: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen 272
Tabelle 4-80: Ausmaß des Zusatznutzens von Lenvatinib vs. Sorafenib auf Endpunktebene 280
Tabelle 4-81: Übersicht zur Gesamtaussage zum Zusatznutzen von Lenvatinib vs. Sorafenib 283
Tabelle 4-82: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 284
Tabelle 4-83: Liste der eingeschlossenen Studien 287
Tabelle 4-84 (Anhang): Studiendesign und -methodik für Studie 304 314
Tabelle 4-85 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie 304 333

Eigene Tabellen

Tabelle 4- A: Übersicht der Ein- und Ausschlusskriterien der Studien, die in die Nutzenbewertung eingehen 28
Tabelle 4-B: Ausmaß des Zusatznutzens von Lenvatinib vs. Sorafenib auf Endpunktebene .. 42
Tabelle 4-C: Übersicht zur Gesamtaussage zum Zusatznutzen von Lenvatinib vs. Sorafenib 45
Tabelle 4-D: Übersicht der Ein- und Ausschlusskriterien der Studien, die in die Nutzenbewertung eingehen 52
Tabelle 4-E: Patientencharakteristika der Studie 304 zu Baseline 60
Tabelle 4-F: Subgruppenanalysen der Studie 304 78

Abbildungsverzeichnis

Seite

Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 85

Abbildung 4-2: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel 108

Abbildung 4-3: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 115

Abbildung 4-4: Sensitivitätsanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (IIR, RECIST v.1.1) aus Studie 304 mit dem zu bewertenden Arzneimittel..... 117

Abbildung 4-5: Sensitivitätsanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (Prüfarzt, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel 118

Abbildung 4-6: Waterfall Plots für die prozentuale Veränderung der aufsummierten Durchmesser der Zielläsionen für die einzelnen Patienten (IIR, mRECIST) 123

Abbildung 4-7: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 127

Abbildung 4-8: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Übelkeit** und **Erbrechen** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 128

Abbildung 4-9: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 128

Abbildung 4-10: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 129

Abbildung 4-11: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 129

Abbildung 4-12: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 130

Abbildung 4-13: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Obstipation** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 130

Abbildung 4-14: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 131

Abbildung 4-15: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 133

Abbildung 4-16: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Übelkeit und Erbrechen** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 134

Abbildung 4-17: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 134

Abbildung 4-18: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 135

Abbildung 4-19: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 135

Abbildung 4-20: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 136

Abbildung 4-21: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Obstipation** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 136

Abbildung 4-22: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 137

Abbildung 4-23: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 141

Abbildung 4-24: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ikterus** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 142

Abbildung 4-25: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 142

Abbildung 4-26: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 143

Abbildung 4-27: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Abdominale Schwellung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 143

Abbildung 4-28: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 145

Abbildung 4-29: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Ikterus** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 146

Abbildung 4-30: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 146

Abbildung 4-31: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 147

Abbildung 4-32: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Abdominale Schwellung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 147

Abbildung 4-33: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 7 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 151

Abbildung 4-34: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 152

Abbildung 4-35: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel 153

Abbildung 4-36: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 157

Abbildung 4-37: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Rollenfunktionsniveau** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 158

Abbildung 4-38: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 158

Abbildung 4-39: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 159

Abbildung 4-40: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 159

Abbildung 4-41: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Finanzielle Schwierigkeiten** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel..... 160

Abbildung 4-42: Hauptanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala

Globaler Gesundheitsstatus des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 160

Abbildung 4-43: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 163

Abbildung 4-44: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Rollenfunktionsniveau** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 163

Abbildung 4-45: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 164

Abbildung 4-46: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 164

Abbildung 4-47: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 165

Abbildung 4-48: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Finanzielle Schwierigkeiten** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 165

Abbildung 4-49 Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel 166

Abbildung 4-50: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 170

Abbildung 4-51: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 171

Abbildung 4-52: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 171

Abbildung 4-53: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 173

Abbildung 4-54: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 174

Abbildung 4-55: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel 174

Abbildung 4-56: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** aus Studie 304 mit dem zu bewertenden Arzneimittel 179

Abbildung 4-57: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **SUE** aus Studie 304 mit dem zu bewertenden Arzneimittel 179

Abbildung 4-58: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE, die zum Therapieabbruch führten** aus Studie 304 mit dem zu bewertenden Arzneimittel 180

Abbildung 4-59: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** aus Studie 304 mit dem zu bewertenden Arzneimittel 180

Abbildung 4-60: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Diarrhö mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 182

Abbildung 4-61: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte Gamma-Glutamyl Transferase mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 182

Abbildung 4-62: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **hepatische Enzephalopathie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 183

Abbildung 4-63: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Fatigue mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 183

Abbildung 4-64: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Aszites mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 184

Abbildung 4-65: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Hyponatriämie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 184

Abbildung 4-66: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Anämie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 185

Abbildung 4-67: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte Alaninaminotransferase mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 185

Abbildung 4-68: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **palmar-plantares Erythrodysästhesiesyndrom mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 186

Abbildung 4-69: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Asthenie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 186

Abbildung 4-70: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erniedrigte Leukozytenzahl mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 187

Abbildung 4-71: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erniedrigte Neutrophilenzahl mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 187

Abbildung 4-72: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte alkalische Phosphatase im Blut mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 188

Abbildung 4-73: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte Lipase mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 188

Abbildung 4-74: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Abdominalschmerz mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel 189

Abbildung 4-75: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **PFS** (IIR, mRECIST) (Region[a] „Asien-Pazifik") aus Studie 304 mit dem zu bewertenden Arzneimittel 195

Abbildung 4-76: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (IIR, mRECIST) (Region[a] „Westliche Regionen") aus Studie 304 mit dem zu bewertenden Arzneimittel 195

Abbildung 4-77: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 (Ätiologie HBV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 198

Abbildung 4-78: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 (Ätiologie HBV „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 199

Abbildung 4-79: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Übelkeit** und **Erbrechen** des EORTC QLQ-C30 (MPVI oder ES oder beides „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 199

Abbildung 4-80: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Übelkeit** und **Erbrechen** des EORTC QLQ-C30 (MPVI oder ES oder beides „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 200

Abbildung 4-81: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 (Ätiologie HBV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 200

Abbildung 4-82: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 (Ätiologie HBV „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 201

Abbildung 4-83: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (MPVI „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 201

Abbildung 4-84: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (MPVI „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 202

Abbildung 4-85: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe**

des EORTC QLQ-C30 (Baseline AFP „< 200 ng/mL“) aus Studie 304 mit dem zu bewertenden Arzneimittel 202

Abbildung 4-86: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (Baseline AFP „≥ 200 ng/mL“) aus Studie 304 mit dem zu bewertenden Arzneimittel 203

Abbildung 4-87: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (MPVI „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 203

Abbildung 4-88: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (MPVI „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 204

Abbildung 4-89: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IRR) „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 204

Abbildung 4-90: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IRR) „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 205

Abbildung 4-91: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 205

Abbildung 4-92: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 206

Abbildung 4-93: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Baseline Altersgruppe „< 65 Jahre) aus Studie 304 mit dem zu bewertenden Arzneimittel 206

Abbildung 4-94: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Baseline Altersgruppe „≥ 65 bis < 75 Jahre”) aus Studie 304 mit dem zu bewertenden Arzneimittel 207

Abbildung 4-95: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Baseline Altersgruppe „≥ 75 Jahre”) aus Studie 304 mit dem zu bewertenden Arzneimittel 207

Abbildung 4-96: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö**

des EORTC QLQ-C30 (Ätiologie HBV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 208

Abbildung 4-97: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HBV „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 208

Abbildung 4-98: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HCV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 209

Abbildung 4-99: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HCV „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 209

Abbildung 4-100: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 (Baseline Altersgruppe „< 65 Jahre") aus Studie 304 mit dem zu bewertenden Arzneimittel 216

Abbildung 4-101: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 (Baseline Altersgruppe „≥ 65 bis < 75 Jahre") aus Studie 304 mit dem zu bewertenden Arzneimittel 216

Abbildung 4-102: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 (Baseline Altersgruppe „≥ 75 Jahre") aus Studie 304 mit dem zu bewertenden Arzneimittel 217

Abbildung 4-103: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 (Baseline AFP „< 200 ng/mL") aus Studie 304 mit dem zu bewertenden Arzneimittel 217

Abbildung 4-104: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 (Baseline AFP „≥ 200 ng/mL") aus Studie 304 mit dem zu bewertenden Arzneimittel 218

Abbildung 4-105: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS (MPVI oder ES oder beides „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 221

Abbildung 4-106: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS (MPVI oder ES oder beides „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel 221

Abbildung 4-107: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel 224

Abbildung 4-108: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 225

Abbildung 4-109: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Region[a] „Asien-Pazifik“) aus Studie 304 mit dem zu bewertenden Arzneimittel 226

Abbildung 4-110: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Region[a] „westliche Regionen“) aus Studie 304 mit dem zu bewertenden Arzneimittel 227

Abbildung 4-111: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 228

Abbildung 4-112: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 228

Abbildung 4-113: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI oder ES oder beides „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 229

Abbildung 4-114: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI oder ES oder beides „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 230

Abbildung 4-115: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 231

Abbildung 4-116: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 231

Abbildung 4-117: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (BCLC Staging „Stadium B“) aus Studie 304 mit dem zu bewertenden Arzneimittel 232

Abbildung 4-118: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (BCLC Staging „Stadium C“) aus Studie 304 mit dem zu bewertenden Arzneimittel 232

Abbildung 4-119: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 (Region[a] „Asien-Pazifik“) aus Studie 304 mit dem zu bewertenden Arzneimittel 233

Abbildung 4-120: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 (Region[a] „Westliche Regionen“) aus Studie 304 mit dem zu bewertenden Arzneimittel 234

Abbildung 4-121: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IIR) „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 235

Abbildung 4-122: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IIR) „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 235

Abbildung 4-123: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Baseline Gewichtsgruppe „< 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel 236

Abbildung 4-124: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Baseline Gewichtsgruppe „≥ 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel 236

Abbildung 4-125: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 237

Abbildung 4-126: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 237

Abbildung 4-127: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 (Baseline AFP < 200 ng/mL) aus Studie 304 mit dem zu bewertenden Arzneimittel 243

Abbildung 4-128: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 (Baseline AFP ≥ 200 ng/mL) aus Studie 304 mit dem zu bewertenden Arzneimittel 244

Abbildung 4-129: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (Geschlecht „Männlich“) aus Studie 304 mit dem zu bewertenden Arzneimittel 244

Abbildung 4-130: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (Geschlecht „Weiblich“) aus Studie 304 mit dem zu bewertenden Arzneimittel 245

Abbildung 4-131: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (MPVI oder ES oder beides „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 246

Abbildung 4-132: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (MPVI oder ES oder beides „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 247

Abbildung 4-133: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 (Leberzirrhose zu Baseline (IIR) „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel 247

Abbildung 4-134: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 (Leberzirrhose zu Baseline (IIR) „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 248

Abbildung 4-135: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel..... 252

Abbildung 4-136: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 252

Abbildung 4-137: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HCV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel..... 253

Abbildung 4-138: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HCV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel 253

Abbildung 4-139: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE, die zum Therapieabbruch führten** (Vorbehandlung mit TACE „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel..... 254

Abbildung 4-140: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE, die zum Therapieabbruch führten** (Vorbehandlung mit TACE „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel..... 254

Abbildung 4-141: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Region[a] „Asien-Pazifik“) aus Studie 304 mit dem zu bewertenden Arzneimittel..... 255

Abbildung 4-142: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Region[a] „Westliche Regionen“) aus Studie 304 mit dem zu bewertenden Arzneimittel 256

Abbildung 4-143: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Baseline Gewichtsgruppe „< 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel 257

Abbildung 4-144: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Baseline Gewichtsgruppe „≥ 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel 257
Abbildung 4-145: Flow Chart Studie 304 331

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AASLD	American Association for the Study of Liver Diseases
ADO	Arbeitsgemeinschaft Dermatologische Onkologie
AFP	Alpha-Fetoprotein
ALP	Alkalische Phosphatase
ALT	Alanin-Aminotransferase
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANC	Absolute Neutrophile Count
AST	Aspartat-Aminotransferase
BCLC	Barcelona Clinic Liver Cancer
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BOR	Best Overall Response
BP	blood pressure
BSC	Best Supportive Care
bzgl.	bezüglich
bzw.	beziehungsweise
CBR	Clinical Benefit Rate
CHMP	Committe for Medicinal Products for Human Use
CI	Confidence Interval
CR	Complete Response
CRF	Case Report Form
CSF	colony-stimulating factor
CSR	Clinical Study Report
CT	Computertomografie
CTCAE	Common Terminology Criteria for Adverse Events
d.h.	das heißt
DCR	Disease Control Rate
DGHO	Deutschen Gesellschaft für Hämatologie und Onkologie
dL	Deziliter
DSMB	Data Safety Monitoring Board
EASL	European Association for the Study of the Liver

ECOG-PS	Eastern Cooperative Oncology Group-Performance Status
EMA	European Medicines Agency
EORTC	European Organization for Research and Treatment of Cancer
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
EORTC QLQ-HCC18	European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire
EQ-5D	EuroQoL Five Dimension Health Questionnaire
ES	Extrahepatic Spread
EU	Europäische Union
EU-CTR	EU Clinical Trials Register
FAS	Full Analysis Set
FDA	Food and Drug Administration
FGF	Fibroblast Growth Factor
FGFR1	FGF Rezeptor
G-BA	Gemeinsamer Bundesausschuss
G-CSF	granulocyte colony-stimulating factor
Hb	Hämoglobin
HBV	Hepatitis B Virus
HCC	hepatocellular carcinoma
hCG	humanes Choriongonadotropin
HCV	Hepatitis C Virus
HIV	Humanes Immundefizienz-Virus
HR	Hazard Ratio
HRQoL	Health-related quality of life
i.d.R.	in der Regel
ICL	Imaging Core Laboratory
ICTRP	International Clinical Trials Registry Platform Search Portal
IIR	Independent Imaging Review
INR	International Normalized Ratio
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IU	international unit
IxRS	interaktives Voice and Web Response System

KI	Konfidenzintervall
KM	Kaplan-Meier
LS	Least Square
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Difference
mL	Milliliter
MPVI	Macroscopic Portal Vein Invasion
mRECIST	modified Response Evaluation Criteria in Solid Tumors
MRI	Magnetic Resonance Imaging
MUGA	multiple gated acquisition scan
NE	not estimable
Nr	Nummer
NYHA	New York Heart Association
OR	Odds Ratio
ORR	Objective Response Rate
OS	Overall Survival
PD	progressive disease
PDGF	Platelet-Derived Growth
PFS	Progression Free Survival
PPS	Per-Protocol-Analysis-Set
PR	Partial Response
PRES	Posteriores Reversibles Enzephalopathiesyndrom
PRO	patient-reported outcome
QoL	Quality of Life
QTc	frequenzkorrigierte QT-Zeit
RD	Risikodifferenz
RECIST	Response Evaluation Criteria in Solid Tumors
RR	Risk Ratio
RTK	Rezeptor-Tyrosinkinase
SAG	Scientific Advisory Group
SAP	statistischer Analyseplan
SAS	Safety Analysis
SD	Standard Deviation

SUE	Schwerwiegendes UE
TACE	Transarterielle Chemoembolisation
TEAE	Treatment Emergent Adverse Events
TTP	Time to Progression
u.a.	under Anderem
UE	unerwünschtes Ereignis
ULN	upper limit of normal
VAS	Visual Analogue Scale
VEGF	Vascular Endothelial Growth Factor
VEGFR1	VEGF Rezeptor 1
VHP	Voluntary Harmonisation Procedure
vs.	versus
z.B.	zum Beispiel
ZVT	zweckmäßige Vergleichstherapie(n)
µg	Mikrogramm
CONSORT	Consolidated Standards of Reporting Trials
DIMDI	Deutsches Institut für Medizinische Dokumentation
ITT	Intention-to-treat
MTC	Mixed Treatment Comparison
RCT	Randomized Controlled Trial
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
WHO	World Health Organization

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

– Zusammenfassung (Abschnitt 4.1)

– Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)

– Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)

– eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)

– ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

Wie ist die quantitative und qualitative Evidenz aus RCT (Randomized Controlled Trial(s), randomisierte(n) kontrollierte(n) Studie(n)) hinsichtlich eines Zusatznutzens von Lenvatinib als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem hepatozellulären Karzinom (HCC), die zuvor noch keine systemische Therapie erhalten haben, im Vergleich zu Sorafenib hinsichtlich der patientenrelevanten Endpunkte Mortalität (Gesamtüberleben), Morbidität und gesundheitsbezogene Lebensqualität sowie der patientenrelevanten Endpunkte zu Sicherheit und Verträglichkeit zu bewerten?

Datenquellen

Der Nachweis des Nutzens und des Zusatznutzens von Lenvatinib erfolgt im Rahmen eines direkten Vergleichs von Lenvatinib gegenüber der systemischen Therapie mit Sorafenib basierend auf den Ergebnissen der randomisierten kontrollierten pivotalen Phase-III-Studie 304.

Die Studie 304 entspricht nach der in der Verfahrensordnung des Gemeinsamen Bundesausschusses (G-BA) dargestellten Evidenzhierarchie der Evidenzstufe Ib (G-BA 2018b). Das Verzerrungspotenzial auf Studienebene ist insgesamt als niedrig einzustufen.

Ein-/Ausschlusskriterien für Studien

Tabelle 4- A: Übersicht der Ein- und Ausschlusskriterien der Studien, die in die Nutzenbewertung eingehen

Kategorie	Einschlusskriterien		Ausschlusskriterien	
Population	E1	▪ Erwachsene Patienten mit fortgeschrittenem oder inoperablem hepatozellulärem Karzinom (HCC), die zuvor noch keine systemische Therapie erhalten haben	A1	▪ Patienten < 18 Jahre, die nicht E1 entsprechen
Intervention	E2	▪ Patienten mit einem Körpergewicht ≥ 60 kg zu Baseline: Lenvatinib (12 mg), oral ▪ Patienten mit einem Körpergewicht < 60 kg zu Baseline: Lenvatinib (8 mg), oral	A2	▪ Lenvatinib in anderer als der gemäß Fachinformation empfohlenen Tagesdosis von Lenvatinib (12 mg für Patienten mit einem Körpergewicht ≥ 60 kg oder 8 mg für Patienten mit einem Körpergewicht < 60 kg)
Vergleichstherapie	E3	▪ Sorafenib in zugelassener Tagesdosis (800 mg), oral	A3	▪ Sorafenib in nicht zugelassener Tagesdosis
Patientenrelevante Endpunkte	E4	Mindestens einer der gelisteten Endpunkte: ▪ OS ▪ PFS ▪ Morbidität (Symptome) ▪ Gesundheitsbezogene Lebensqualität ▪ Unerwünschte Ereignisse	A4	▪ Andere Endpunkte als in E4 ▪ Punkt- und Streumaße nicht vorhanden
Studientypen	E5	▪ RCT	A5	▪ Tierexperimentelle Studien ▪ Keine RCT
Studiendauer	E6	▪ Keine Einschränkung bezüglich der Studiendauer	A6	▪ Keine Einschränkung bezüglich der Studiendauer
Sprache	E7	▪ Englisch oder Deutsch	A7	▪ Andere als Englisch oder Deutsch
Publikationstyp	E8	▪ Vollpublikation zu RCT	A8	▪ Mehrfachpublikationen ohne relevante Zusatzinformationen ▪ Kein Studienbericht
Studienstatus	E9	▪ Studien, für die Ergebnisse vorliegen	A9	▪ Noch nicht begonnene, laufende oder abgeschlossene Studien, für die noch keine Ergebnisse vorliegen

kg: Kilogramm; mg: Milligramm; OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Die Bewertung erfolgte auf Studienebene und separat für jeden patientenrelevanten Endpunkt auf Basis des klinischen Studienberichts. Die Extraktion und Dokumentation wurde unter Verwendung von Anhang 4-F und mittels der folgenden endpunktübergreifenden und endpunktspezifischen Aspekte durchgeführt.

Die Beurteilung des Verzerrungspotenzials, das aus der Selektion der Studienpopulation resultiert, wurde anhand der Randomisierung der Patienten vorgenommen. Dabei mussten die Erzeugung der Randomisierungssequenz und die daraus resultierende Verteilung der Patienten adäquat durchgeführt und dokumentiert worden sein. Ebenfalls musste die Randomisierungssequenz den Behandlern verborgen geblieben sein. Um eine Verzerrung der Studienergebnisse durch den Patienten oder den Behandler zu vermeiden, sollten beide in adäquater Weise verblindet worden sein. Bei offen durchgeführten Studien ist die mögliche Auswirkung auf das Verzerrungspotenzial zu diskutieren. Des Weiteren wurde kritisch überprüft, ob die gültigen Kriterien für das Studiendesign eingehalten wurden.

Um eine Verzerrung auf Ebene der Studienendpunkte ausschließen zu können, mussten alle festgelegten Zielgrößen vollständig berichtet werden. Dabei wird die ITT-Population als die Studienpopulation mit dem geringsten Verzerrungspotenzial betrachtet. Die Einhaltung des ITT-Prinzips wurde im Einklang mit den Vorgaben des vorliegenden Dokuments u.a. dann als adäquat umgesetzt bewertet, wenn der Nichtberücksichtigungsanteil (von z.B. Protokollverletzern und Lost-to-Follow-up-Patienten) in der Auswertung < 5% betrug.

Eine zusammenfassende Bewertung des Verzerrungspotenzials auf Studien- und Endpunktebene erfolgte in den Kategorien der Ergebnissicherheit „hoch“ und „niedrig“. Ein niedriges Verzerrungspotenzial lag dann vor, wenn davon ausgegangen werden konnte, dass die Ergebnisse nicht relevant verzerrt waren. Eine relevante Verzerrung ist dann anzunehmen, wenn sich die Ergebnisse insgesamt oder die Ergebnisse auf Endpunktebene grundlegend verändern.

Wurde das Verzerrungspotenzial auf Studienebene (A) mit „hoch“ bewertet, erfolgte die Einstufung des Verzerrungspotenzials auf Endpunktebene (B) ebenfalls als „hoch“. Ein hohes Verzerrungspotenzial führte nicht zum Ausschluss der Daten.

Das Ausmaß der qualitativen Ergebnissicherheit auf Studien- und Endpunktebene wurde anschließend auf Basis des Verzerrungspotenzials und der European Medicines Agency (Europäische Arzneimittel Agentur, EMA)-Kriterien für die Zulassung auf Grundlage einer einzelnen pivotalen Studie bestimmt (EMA 2001). Die letztgenannten Kriterien werden durch die Erteilung einer Zulassung für Lenvima® zur Behandlung des HCC auf Basis einer Phase-III-Studie als gegeben betrachtet.

Patientenrelevante Endpunkte

Im vorliegenden Nutzendossier werden die folgenden patientenrelevanten Endpunkte ausgewertet:

Mortalität

- OS

Morbidität

- PFS
- EORTC QLQ-C30 (Symptome)
    - Fatigue
    - Übelkeit und Erbrechen
    - Schmerz
    - Dyspnoe
    - Schlaflosigkeit
    - Appetitminderung
    - Obstipation
    - Diarrhö
- EORTC QLQ-HCC18 (Symptome)
    - Fatigue
    - Ikterus
    - Schmerz
    - Fieber
    - Abdominale Schwellung
- EQ-5D VAS

Gesundheitsbezogene Lebensqualität

- EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität)
    - Physische Funktion
    - Rollenfunktionsniveau
    - Emotionale Funktion
    - Kognitive Funktion
    - Soziale Funktion
    - Globaler Gesundheitsstatus
    - Finanzielle Schwierigkeiten
- EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität)

- Körperbild
- Ernährung
- Sexualleben

Sicherheit und Verträglichkeit

- UE
- SUE
- UE, die zum Therapieabbruch führten
- UE mit CTCAE Grad ≥ 3

Die Objektive Ansprechrate (engl. Objective Response Rate, ORR) wird in diesem Dossier als ergänzende Analyse berichtet, aber nicht zur Ableitung des Zusatznutzens herangezogen.

Mortalität

Gesamtüberleben (Overall Survival, OS)

OS wurde definiert als die Zeitspanne zwischen dem Zeitpunkt der Randomisierung bis zum Eintreten des Todes aufgrund jeglicher Ursache. Lost-to-follow-up-Patienten wurden zum letzten Zeitpunkt, an dem sie bekanntermaßen am Leben waren, zensiert. Überlebende Patienten wurden zum Zeitpunkt des Datenschnitts zensiert (Eisai Inc. 2017b).

Gemäß des Rapid Reports des IQWiG stellt das OS ein herausragendes Therapieziel in der Behandlung onkologischer Erkrankungen dar. Entsprechend ist der Endpunkt OS als unmittelbar patientenrelevant anzusehen (IQWiG 2011).

Morbidität

Progressionsfreies Überleben (Progression-free survival, PFS)

PFS ist definiert als Zeitspanne zwischen dem Zeitpunkt der Randomisierung bis zum Zeitpunkt einer dokumentierten Krankheitsprogression oder bis zum Zeitpunkt des Todes aufgrund jeglicher Ursache, abhängig davon, welches dieser beiden Ereignisse zuerst eintrat (Eisai Inc. 2017b). Somit dokumentiert PFS den direkten Effekt einer Behandlung auf die Tumorprogression unter Verwendung objektiver und quantitativer Methodik und ist als wertvoller Endpunkt für klinische Studien mit krankheitsmodifizierend wirkenden Substanzen anerkannt (FDA 2007). Der Zeitpunkt der objektiven Krankheitsprogression war in der Studie 304 definiert als der früheste Zeitpunkt einer radiologisch bestätigten Krankheitsprogression. Patienten, die die Therapie aufgrund einer Krankheitsprogression beendeten oder bei denen keine radiologische Bestätigung für die Krankheitsprogression vorlag, wurden zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert. Patienten, die eine neue Therapie anfingen oder die nach mehr als einer verpassten Visite oder nach bis zu 28 Tagen der letzten Dosis der Studienmedikation verstarben oder bei denen eine Progression auftrat, wurden ebenfalls zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert (Eisai Inc. 2016).

Die Beurteilung der Tumorprogression erfolgte in der vorliegenden Studie 304 sowohl mittels verblindeter und unabhängiger Bewertung radiologischer Verfahren (engl. Independent Imaging Review, IIR) gemäß den speziell für das HCC entwickelten modified Response Evaluation Criteria in Solid Tumors (mRECIST)-Kriterien sowie den Response Evaluation Criteria in Solid Tumors (RECIST)-Kriterien v.1.1. Desweiteren erfolgte eine Beurteilung durch den Prüfarzt gemäß den mRECIST-Kriterien.

Bei den RECIST-Kriterien handelt es sich um eine etablierte Methodik, um über die Größenänderung der Zielläsionen die Tumorantwort auf die Therapie zu bewerten. Dabei wird die Größe der Zielläsionen mit einer einfachen linearen Summation gemessen, wobei nicht zwischen nekrotischem und lebensfähigem Gewebe unterschieden wird. In mehreren HCC-Studien wurde allerdings nur eine schwache Korrelation zwischen dem klinischen Nutzen neuer systemischer Therapien wie Sorafenib oder lokoregionärer Therapien und der konventionellen radiologischen Messung der Tumorantwort nachgewiesen (Lencioni und Llovet 2010). Limitationen der Messung der Tumorantwort mithilfe der RECIST-Kriterien ergeben sich aus (1) dem möglicherweise fehlenden Umbau von zirrhotischem Lebergewebe um den Tumor herum, (2) der diffusen, infiltrierenden Beschaffenheit vom HCC in vielen zirrhotischen Lebern, (3) der Veränderung der Tumor-Vaskularität, aber nicht der Tumorgröße und (4) der verstärkten Darstellung prämaligner dysplastischer Knoten während der arteriellen Phase und sich daraus ergebender falsch-positiver Bestimmung eines Tumorwachstums (Thomas et al. 2010). Wie bereits im Jahr 2000 von der European Association for the Study of the Liver (EASL) festgestellt wurde, spiegelt sich beim HCC eine umfangreiche Tumor-Nekrose nicht zwangsläufig in einer Reduktion des Umfangs der Läsion wieder (Bruix et al. 2001). Daher wurde von einem EASL-Expertengremium das Konzept der lebensfähigen Zielläsionen, die mittels kontrastverstärkender radiologischer Bildgebung identifiziert werden, vorgeschlagen (Bruix et al. 2001). Dieses Konzept der lebensfähigen Zielläsionen wurde auch in die mRECIST-Kriterien mit aufgenommen (Lencioni und Llovet 2010). Bei der mRECIST-Methodik wird nekrotisches Tumorgewebe durch die fehlende Kontrastbildung innerhalb des Tumors bei der Bildgebung identifiziert.

Da die Veränderungen der intrahepatischen Läsionen durch mRECIST über die Messung der nur lebensfähigen Bereiche der Läsionen besser erfasst werden als über RECIST und die damit einhergehende Messung der gesamten Läsion, wurden in der Studie 304 mRECIST-Kriterien für die Auswertung durch den Prüfarzt herangezogen. Diese Entscheidung erfolgte in Abstimmung mit dem Steering Committee und nach Zustimmung durch die regulatorischen Behörden sowie unter Betrachtung der wissenschaftlichen Evidenz zum Zeitpunkt der Studienplanung (Eisai Inc. 2017b). Für den IIR wurden sowohl die RECIST als auch die mRECIST Kriterien herangezogen.

Gemäß mRECIST ist eine Krankheitsprogression definiert als eine mindestens 20%-ige Zunahme der Summe der Durchmesser der lebensfähigen Zielläsionen, wobei die kleinste dokumentierte Summe der lebensfähigen Zielläsionen seit Behandlungsbeginn als Referenz dient. Eine Zielläsion gemäß mRECIST ist dabei definiert – wie gemäß RECIST – als eine messbare Läsion mit mindestens 1 cm Durchmesser in mindestens einer Dimension (Lencioni

und Llovet 2010). Darüber hinaus sollte die Läsion für wiederholte Messungen geeignet sein und eine intratumorale arterielle Erweiterung bei kontrastverstärkender Computertomografie (CT) oder Magnetic Resonance Imaging (MRI) zeigen (Lencioni und Llovet 2010). Bei Anwendung der mRECIST Kriterien werden hepatische Läsionen gemäß mRECIST und nicht-hepatische Läsionen gemäß den RECIST v.1.1 Kriterien beurteilt. Im Gegensatz zu mRECIST geht gemäß RECIST v.1.1 die komplette Läsion in die Bildung der Summe der Durchmesser mit ein (Lencioni und Llovet 2010).

Die in der Studie 304 herangezogenen mRECIST und RECIST-Kriterien stellen einen anerkannten und etablierten Standard zur Bestimmung des Tumoransprechens beim HCC dar (Lencioni und Llovet 2010).

Kommt es zu einer Tumorprogression, zu Metastasen oder zu einem Rezidiv, ist dies für den betroffenen Patienten schwerwiegend (Berg et al. 2011). Es kommt dabei zur Verschlechterung bzw. zum Auftreten neuer krankheitsspezifischer Symptome sowie zu einer starken psychischen Belastung, z.B. der sog. Progredienzangst (Herschbach et al. 2001; Herschbach und Dinkel 2014; Mehnert et al. 2006). Hierbei handelt es sich um eine reaktive und bewusst wahrgenommene Furcht, die im Rahmen einer schweren, potenziell lebensbedrohlichen, oder zur Behinderung führenden Erkrankung und ihrer Behandlung entsteht (Hesse et al. 2016). Diese Progredienzangst und deren mögliche Behandlung spielt eine immer bedeutendere Rolle im Rahmen einer palliativen Versorgung von Patientinnen und Patienten (Hesse et al. 2016).

Eine Verlängerung des PFS führt zu einem verzögerten Eintreten der Progredienzangst und der krankheitsspezifischen Symptome, was wiederum eine Steigerung der Lebensqualität durch einen positiven Einfluss auf das physische und psychische Wohlbefinden der Patienten mit sich bringt. Das Hinauszögern einer Progression der Erkrankung wird daher von den Patienten als direkter Vorteil wahrgenommen und stellt für den Patienten einen selbst erlebbaren Therapieerfolg dar.

Nicht zuletzt deswegen wurde PFS in den letzten Jahren als Endpunkt von der EMA und der Food and Drug Administration (FDA) bei sehr unterschiedlichen onkologischen Arzneimitteln und Therapieindikationen im Rahmen klinischer Studien als patientenrelevanter Endpunkt für die Zulassung neuer Arzneimittel akzeptiert und bestätigt. Auch das Bundesinstitut für Arzneimittel (BfArM) erachtet das PFS als klinisch relevanten Endpunkt (BfArM 2011). Ein länger anhaltendes PFS wird dabei als relevanter Endpunkt von Patientennutzen anerkannt, wenn das Ausmaß des Behandlungseffektes ausreichend groß ist, um etwaige Verträglichkeitsprobleme zu überwiegen (EMA 2012). Das Committe for Medicinal Products for Human Use (CHMP) der EMA forderte demnach bereits im Jahr 2012 in einer Stellungnahme zur Überarbeitung der „Anticancer Guideline“, dass die alleinstehende klinische Relevanz von PFS anerkannt wird. Die CHMP Scientific Advisory Group (SAG) for Oncology, in der ausgewählte europäische Experten Änderungen der Anticancer Guideline berieten, stellten fest, dass PFS als klinisch relevanter Endpunkt zu betrachten ist, selbst wenn Verbesserungen im PFS nicht mit einer dokumentierten Verlängerung des OS einhergingen, solange es keine Hinweise auf negative Auswirkungen auf das Gesamtüberleben gibt. Dies wurde hauptsächlich begründet mit der Wichtigkeit, die das Hinauszögern des erstmaligen

Auftretens oder die Verschlechterung von Symptomen in Verbindung mit einer Progression sowie einer nächsten Therapielinie für den Patienten besitzt (EMA 2012). Somit ist PFS explizit nicht als Surrogat für OS, sondern als alleinstehender, patientenrelevanter Endpunkt zu betrachten.

Auch diverse Fachgesellschaften teilen diese Meinung. Die Arbeitsgemeinschaft Dermatologische Onkologie (ADO) definiert ein länger anhaltendes mehrmonatiges PFS und den damit einhergehenden Abstand zu einer lebensbedrohlichen Situation als patientenrelevanten Endpunkt. Eine Zeit ohne Progression und damit eine Zeit ohne Todesängste sieht die ADO eindeutig als relevantes Ziel für den Patienten (ADO 2012). Nach Einschätzung der Deutschen Gesellschaft für Hämatologie und Onkologie (DGHO) ist das PFS ein sinnvoller und gebräuchlicher Endpunkt bei allen Therapieinterventionen mit palliativer Zielsetzung. Gegenüber dem Endpunkt Überlebenszeit hat der Endpunkt PFS den Vorteil einer geringeren Beeinflussung durch spätere Störfaktoren wie Cross-over und / oder Folgetherapien (DGHO 2013).

Ergänzend zu PFS wird in dem vorliegenden Dossier die ORR berichtet. Die ORR wurde definiert als der Anteil der Patienten mit dem besten Gesamtansprechen (engl. Best Overall Response, BOR) in Form eines vollständigen Tumoransprechens (engl. Complete Response, CR) oder eines teilweisen Tumoransprechens (engl. Partial Response, PR). Die ORR stellt somit ein Maß für die Kontrolle des Tumorwachstums dar (Eisai Inc. 2017b).

Ein Ansprechen auf die Therapie wurde, analog zum PFS, unter Verwendung von IIR gemäß mRECIST bzw. als Sensitivitätsanalyse gemäß RECIST v.1.1 sowie durch den Prüfarzt gemäß mRECIST gemessen. Unter CR versteht man dabei gemäß mRECIST den vollständigen Rückgang aller intratumoralen arteriellen Erweiterungen der Zielläsionen (Lencioni und Llovet 2010) und unter PR eine Reduktion der Summe der Durchmesser der lebensfähigen Zielläsionen um mindestens 30% im Vergleich zum Ausgangswert (Lencioni und Llovet 2010). Gemäß RECIST v.1.1 ist CR definiert als der vollständige Rückgang aller Zielläsionen. PR ist hingegen definiert als die Reduktion der Summe der Durchmesser der Zielläsionen um mindestens 30% im Vergleich zum Ausgangswert (Lencioni und Llovet 2010).

Die ORR wird in diesem Dossier als ergänzende Analyse berichtet, aber nicht zur Ableitung des Zusatznutzens herangezogen.

EORTC QLQ-C30

Der Fragebogen EORTC QLQ-C30 ist in der Onkologie ein häufig eingesetztes Instrument zur Bewertung der Morbidität und Lebensqualität mit insgesamt 30 Fragen. Die gestellten Fragen beziehen sich dabei auf die Symptome Fatigue, Übelkeit und Erbrechen, Schmerz, Dyspnoe, Schlaflosigkeit, Appetitminderung, Obstipation, Diarrhö sowie auf die Funktionen Physische Funktion, Rollenfunktionsniveau, Emotionale Funktion, Kognitive Funktion, Soziale Funktion, Finanzielle Schwierigkeiten und allgemein auf den Globalen Gesundheitsstatus. Die einzelnen Items werden dabei auf einer Skala von 1 bis 4 bzw. 1 bis 7, für die Fragen zum globalen Gesundheitsstatus, bewertet, die zur Auswertung in eine Skala von 0 bis 100 transformiert wird. Höhere Werte auf den Funktionsskalen und der Globalskala bedeuten eine höhere

Lebensqualität und auf den Symptomskalen eine Verschlechterung der Symptome (Giesinger et al. 2016). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen (engl. minimal important difference, MID) (Cocks et al. 2011; Osoba et al. 1998).

Der EORTC QLQ-C30 ist eines der höchstentwickelten und am häufigsten verwendeten Messinstrumente zur Erhebung der Lebensqualität bei Patienten mit einer Krebserkrankung. Die Validität des Patientenfragebogens wurde für verschiedene onkologische Indikationen bestätigt (Aaronson et al. 1993). Darüber hinaus ist der EORTC QLQ-C30 als validiertes Messinstrument gemäß Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) anerkannt (G-BA 2014; IQWiG 2013, 2014; Medac 2013).

Im Rahmen der vorliegenden Bewertung werden die Funktionsskalen zur Erfassung der gesundheitsbezogenen Lebensqualität herangezogen, während die Symptomskalen der Messung der Krankheitssymptome dienen.

EORTC QLQ-HCC18

Der EORTC QLQ-HCC18 ist ein HCC-spezifischer Fragebogen mit 18 Fragen bezüglich der Symptome Fatigue, Ikterus, Schmerz, Fieber und Abdominale Schwellung sowie der Items Körperbild, Ernährung und Sexualleben und wurde entwickelt, um die Sensitivität und Spezifität von HCC-bezogenen Fragestellungen im Hinblick auf die gesundheitsbezogene Lebensqualität zu erhöhen. Die einzelnen Items werden auf einer Skala von 1 bis 4 bewertet, welche wiederum linear in eine Skala von 0 bis 100 umgewandelt wird. Der Wert 100 stellt dabei den schlechtesten Status dar (Chie et al. 2012). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen.

Für die Bewertung der Morbidität wurden die Symptomskalen zu Fatigue, Ikterus, Schmerz, Fieber, Abdominale Schwellung und für die Bewertung der gesundheitsbezogenen Lebensqualität die Items Körperbild, Ernährung und Sexualleben des EORTC QLQ-HCC18 herangezogen.

Sowohl die Ergebnisse des EORTC QLQ-C30 als auch die Ergebnisse des EORTC QLQ-HCC18 sind hochrelevante prognostische Faktoren für das Gesamtüberleben bei neu diagnostizierten HCC-Patienten (Li et al. 2017). Das hierbei verwendetet Scoring-Verfahren ist eine effektive Möglichkeit zur Zusammenfassung, Analyse und Interpretation von Rohdaten zur Morbidität und gesundheitsbezogenen Lebensqualität und macht die beiden Fragebögen EORTC QLQ-C30 und EORTC QLQ-HCC18 in der klinischen Praxis aussagekräftig und kommunikationsfähig (Chie et al. 2012; Li et al. 2017).

EQ-5D VAS

Der Fragebogen EQ-5D ist ein von der internationalen Forschungsgruppe EuroQoL entwickeltes, verlässliches und validiertes Messinstrument für den Gesundheitszustand eines Patienten. Er besteht aus einem deskriptiven Teil und einer visuellen Analogskala (engl. Visual Analogue Scale, VAS), die zusätzlich in den Fragebogen aufgenommen wurde, um den

aktuellen Gesundheitszustand auf einer Skala von 0 bis 100 zu messen, wobei 0 für den "schlechtesten denkbaren Gesundheitszustand" und 100 für den "besten denkbaren Gesundheitszustand" steht. (Pickard et al. 2012). Der Fragebogen wurde und wird derzeit in zahlreichen Studien mit Krebspatienten sowie speziell bei Patienten mit HCC eingesetzt (Pickard et al. 2012; Teckle et al. 2013). Eine Veränderung auf der VAS eines Patienten um mindestens 7 oder 10 mm im Vergleich zum Ausgangswert wurde hierbei als Response definiert (Pickard et al. 2007).

Die EQ-5D VAS stellt somit ein allgemeines quantitatives Maß für den Gesundheitszustand eines Patienten entsprechend seiner eigenen Beurteilung dar. Im Sinne der Berücksichtigung der Lebensqualität und Symptomatik der Patienten im Rahmen der Nutzenbewertung ist die EQ-5D VAS als patientenrelevanter Endpunkt anzusehen (IQWiG 2017).

Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität ist eine wichtige Zielgröße bei der medizinischen Versorgung krebskranker Patienten (IQWiG 2011). Sie spiegelt direkt die Auswirkungen der Erkrankung auf das Befinden des Patienten und seinen Alltag wider und ist damit unmittelbar patientenrelevant. Auch der Erhalt der Lebensqualität stellt gerade in der palliativen Situation, die auf Dauer fast immer mit einer Verschlechterung der Lebensqualität einhergeht, einen patientenrelevanten Endpunkt dar (IQWiG 2017).

Die gesundheitsbezogene Lebensqualität wurde in der eingeschlossenen Studie 304 anhand der Funktionsskalen Physische Funktion, Rollenfunktionsniveau, Emotionale Funktion, Kognitive Funktion, Soziale Funktion, Finanzielle Schwierigkeiten und Globaler Gesundheitsstatus des Fragebogens EORTC QLQ-C30 beurteilt. Außerdem wurden für die Bewertung der gesundheitsbezogenen Lebensqualität die Items Körperbild, Ernährung und Sexualleben des HCC-spezifischen EORTC QLQ-HCC18 herangezogen.

Sicherheit und Verträglichkeit

Im vorliegenden Dossier werden die Kategorien UE, SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 ausführlich dargestellt. UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, werden ergänzend dargestellt.

Die Kodierung der UE erfolgte unter Verwendung des Medical Dictionary for Regulatory Activities (MedDRA) Version 19.1. Die Einteilung in verschiedene Grade erfolgte nach CTCAE Version 4.0.

Alle UE und SUE wurden als Ereignisse erfasst, die während der Behandlung auftraten (engl. Treatment Emergent Adverse Events, TEAE) erfasst. TEAE wurden definiert als jedes Ereignis,

- das zu Baseline nicht präsent war und während oder bis zu 30 Tage nach der Behandlung mit der Studienmedikation auftrat.

- das bei einer Vorbehandlung aufgetreten ist, aber zu Baseline nicht mehr präsent war und während oder bis zu 30 Tage nach der Behandlung mit der Studienmedikation wieder auftrat.
- dessen Schweregrad sich während oder bis zu 30 Tage nach der Behandlung im Vergleich zum Zustand vor der Behandlung mit der Studienmedikation verschlechterte.

Die im Rahmen dieser Nutzenbewertung berichteten UE zählen zu den Aspekten der Sicherheit und Verträglichkeit und sind somit vom IQWiG anerkannte patientenrelevante Endpunkte (IQWiG 2017).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Zur Bewertung des medizinischen Zusatznutzens von Lenvatinib als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben, werden die Ergebnisse einer multizentrischen, randomisierten, aktivkontrollierten und Open-Label durchgeführten Nicht-Unterlegenheits- Phase-III-Studie, der Studie 304, zu den Endpunkten zu Mortalität, Morbidität, gesundheitsbezogene Lebensqualität sowie Sicherheit und Verträglichkeit herangezogen. Das Verzerrungspotenzial der Studie 304 ist insgesamt als niedrig einzustufen.

Mortalität

Für den Wirksamkeitsendpunkt OS zeigte sich in der betrachteten Studienpopulation bezüglich des HR ein numerischer Vorteil für Lenvatinib gegenüber Sorafenib basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell (HR=0,92; 95%-KI [0,79; 1,06]; p=0,2526). Da die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag, konnte hiermit die Nicht-Unterlegenheit von Lenvatinib gegenüber Sorafenib belegt werden, was einen Überlebensvorteil von Lenvatinib gegenüber Placebo impliziert. Die mediane Überlebensdauer unter Lenvatinib betrug 13,6 Monate (95%-KI [12,1; 14,9]), unter Sorafenib betrug sie 12,3 Monate (95%-KI [10,4; 13,9]). Damit ergibt sich ein absoluter Vorteil im medianen Überleben von 1,3 Monaten zugunsten von Lenvatinib.

Die Studie 304 ist somit die erste positive, globale Phase-III-Studie in der Erstlinientherapie des HCC in den letzten 10 Jahren. Darüber hinaus ist die Studie 304 die erste positive Studie im HCC mit einer aktiven Kontrollgruppe. Seit der Zulassung von Sorafenib gab es vier Phase-III-Studien mit zielgerichteten Wirkstoffen, die keine Überlegenheit oder Nicht-Unterlegenheit in der Erstlinientherapie zeigen konnten (Kudo et al. 2018).

Die Ergebnisse der Sensitivitätsanalysen, die auf uni- und multivariater Adjustierung basierten, zeigten konsistent gleichgerichtete Effekte wie die Ergebnisse der Hauptanalyse und bestätigen somit deren Robustheit. In Bezug auf die AFP adjustierte und multivariate adjustierte Analyse zeigte Lenvatinib im Vergleich zu Sorafenib sogar eine statistisch signifikante Überlegenheit im OS. Die in den adjustierten Analysen nachgewiesene statistische Signfikanz kann auf die Imbalance in den Behandlungsarmen bezüglich des für das HCC prognostischen Faktors AFP zurückgeführt werden.

Zusammenfassend liegen für Lenvatinib gegenüber Sorafenib eine Nicht-Unterlegenheit und ein numerischer Vorteil bezüglich der Hauptanalyse vor. Dieser numerische Vorteil im OS wird nach Adjustierung für prognostische Faktoren statistisch signifikant und weist damit auf eine Überlegenheit im OS hin.

Morbidität

In der Hauptanalyse des Endpunkts PFS auf Basis des IIR per mRECIST zeigte sich im HR eine statistisch signifikante und klinisch relevante Überlegenheit bezüglich der Risikoreduktion einer Progression von Lenvatinib gegenüber Sorafenib basierend auf dem Cox-Proportional-Hazard-Modell (HR 0,64; 95%-KI [0,55; 0,75]; p<0,0001). Unter der Behandlung mit Lenvatinib wurde somit eine 36%-ige Verringerung des Progressionsrisikos im Vergleich zur Behandlung mit Sorafenib erreicht. In Bezug auf die mediane Zeit bis zur Krankheitsprogression war das mediane PFS im Lenvatinibarm mit 7,3 Monaten (95%-KI [5,6; 7,5]) mehr als doppelt so lang wie im Sorafenibarm (3,6 Monaten (95%-KI [3,6; 3,7])). Somit ergab sich ein klinisch relevanter absoluter Vorteil von 3,7 Monaten für das mediane PFS zugunsten von Lenvatinib. Die Überlegenheit von Lenvatinib beim PFS konnte dabei unabhängig von der verwendeten Analysemethode bestätigt werden, wie durch die Ergebnisse der zwei durchgeführten Sensitivitätsanalysen gemäß IIR per RECIST v1.1 (HR 0,65; 95%-KI [0,56; 0,77]; p<0,0001) und Prüfarzt per mRECIST (HR 0,66; 95%-KI [0,57; 0,77]; p<0,0001) gezeigt wurde.

Die Ergebnisse der Analysen für den Endpunkt ORR zeigen sowohl in der Tumorbewertung gemäß IIR per mRECIST (OR=5,01; 95%-KI [3,59; 7,01]; p<0,0001) und RECIST v.1.1 (OR=3,34; 95%-KI [2,17; 5,14]; p<0,0001) als auch in der Tumorbewertung durch den Prüfarzt per mRECIST (OR=3,13; 95%-KI [2,15; 4,56]; p<0,0001) eine statistisch signifikante und klinisch relevante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die Zahl der Patienten, die auf eine Therapie mit Lenvatinib bezogen auf das Tumorwachstum ansprachen, war somit um mehr als das Dreifache (gemäß IIR per RECIST 1.1 und Prüfarzt per mRECIST) bzw. sogar mehr als das Fünffache (gemäß IIR per mRECIST) gegenüber Sorafenib erhöht. Legt man die Ergebnisse der unabhängigen radiologischen Untersuchung auf Grundlage der mRECIST Kriterien zugrunde, so heißt dies konkret, dass unter der Behandlung mit Lenvatinib mehr als jeder dritte Patient (40.6%) von der Therapie profitiert, verglichen mit weniger als jedem achten Patienten (12,4%) unter der Behandlung mit Sorafenib. Zudem wurde bei 10 Patienten im Lenvatinibarm eine Complete Response (per mRECIST IIR) – das heißt ein Verschwinden aller Läsionen – beobachtet verglichen mit 4 Patienten im Sorafenibarm (Eisai Inc. 2017a). Die Ergebnisse für den Endpunkt ORR stützen damit die Ergebnisse des Endpunkts PFS.

Die Analyse der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur Morbidität des Fragebogens EORTC QLQ-C30 zeigte für die den Patienten beeinträchtigenden und belastenden Symptome Schmerz (HR=0,81; 95%-KI [0,690; 0,949]; p=0,0091) und Diarrhö (HR=0,52; 95%-KI [0,440; 0,623]; p<0,0001) eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf der Skala Schmerz betrug dabei 1,9 Monate (95%-KI [1,87; 2,76]) für Lenvatinib im Vergleich zu 1,8 Monate

(95%-KI [1,12; 1,87]) für Sorafenib. Bezüglich der Skala Diarrhö betrug der absolute Vorteil 1,9 Monate in der medianen Zeit bis zur ersten Verschlechterung um 10 Punkte zugunsten von Lenvatinib (4,6 Monate; 95%-KI [3,71; 5,52]) im Vergleich zu Sorafenib (2,7 Monate; 95%-KI [1,87; 2,79]). Hinsichtlich der Symptome Fatigue, Übelkeit und Erbrechen, Dyspnoe, Schlaflosigkeit, Appetitminderung und Obstipation zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Auch im Hinblick auf die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur Morbidität des Fragebogens EORTC QLQ-HCC18 konnten die Patienten unter der Behandlung mit Lenvatinib für alle Symptome vergleichbare Therapieerfolge erzielen wie unter der Behandlung mit Sorafenib.

Die Analyse des Fragebogens EQ-5D bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS zeigte für die Patienten unter der Behandlung mit Lenvatinib ebenfalls vergleichbare Therapieerfolge wie unter Sorafenib.

Zusammenfassend besteht für Lenvatinib gegenüber Sorafenib eine statistisch signifikante und klinisch relevante Überlegenheit im PFS und ORR. Darüber hinaus konnte bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Symptome, gemessen auf den Skalen zur Morbidität des EORTC QLQ-C30 und -HCC18, eine statistisch signifikante Überlegenheit bei einzelnen Symptomen sowie ein vergleichbarer Therapieerfolg bei allen weiteren Symptomen von Lenvatinib gegenüber Sorafenib gezeigt werden. Auch bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS konnten vergleichbare Therapieerfolge erzielt werden.

Gesundheitsbezogene Lebensqualität

Bei der Analyse des Rollenfunktionsniveaus des Fragebogens EORTC QLQ-C30 zeigte sich bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf der entsprechenden Skala eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,82; 95%-KI [0,696; 0,962]; p=0,0148). Die mediane Zeit betrug dabei 2,0 Monate (95 5-KI [1,87; 2,83]) für Lenvatinib im Vergleich zu 1,9 Monate (95%-KI [1,84; 1,97]) für Sorafenib. Hinsichtlich der physischen, emotionalen, kognitiven, sozialen und finanziellen Funktion und des globalen Gesundheitsstatus zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Auch die Analysen des Fragenbogens EORTC QLQ-HCC18 bezüglich Körperbild (HR=0,79; 95%-KI [0,671; 0,932]; p=0,0051) und Ernährung (HR=0,80; 95%-KI [0,674; 0,949]; p=0,0102) zeigten eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den entsprechenden Skalen. Bezüglich des Items Körperbild betrug die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte 2,8 Monate (95%-KI [2,73; 3,68]) für Lenvatinib im Vergleich zu 1,9 Monate (95 5-KI [1,87; 2,73]) für Sorafenib. Bezüglich des Items Ernährung betrug die Zeit 4,3 Monate (95%-KI [3,68; 5,52]) für Lenvatinib im Vergleich zu 2,8 Monate (95%-KI [2,07; 3,02]) für Sorafenib.

Zusammenfassend zeigte Lenvatinib gegenüber Sorafenib eine statistisch signifikante Überlegenheit bei drei Endpunkten der gesundheitsbezogenen Lebensqualität (gemessen als Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur gesundheitsbezogenen Lebensqualität des EORTC QLQ-C30 und -HCC18). Für die anderen Endpunkte wurde ein vergleichbarer Therapieerfolg nachgewiesen.

Sicherheit und Verträglichkeit

Da die Gesamtbeobachtungszeiten im Behandlungsarm Lenvatinib um mehr als 75 Tage länger waren als in der Kontrollgruppe, wären die Inzidenzraten und die darauf basierenden Effektmaße RR, OR und RD zuungunsten von Lenvatinib verzerrt, so dass nur sehr eingeschränkt Aussagen zu einem größeren oder vergleichbaren bzw. geringeren Schaden gemacht werden könnten. Aus diesem Grund basiert die Bewertung der Sicherheit und Verträglichkeit von Lenvatinib im vorliegenden Dossier allein auf den Analysen der Zeit bis zum Auftreten des ersten Ereignisses.

Die Ergebnisse zur Analyse der Sicherheit und Verträglichkeit ergaben für die Kategorien SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 weder einen geringeren noch größeren Schaden für Lenvatinib gegenüber Sorafenib. Bezüglich der Kategorie UE zeigte sich bei der Analyse eine signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,84; 95%-KI [0,74; 0,95]).

Ergänzend wurden die Ergebnisse der Analysen der UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, dargestellt. Alle UE mit CTCAE Grad 3 oder 4, mit Ausnahme der UE hepatische Enzephalopathie und palmar-plantares Erythrodysästhesiesyndrom, ergaben weder einen geringeren noch größeren Schaden für Lenvatinib gegenüber Sorafenib. In Bezug auf die Zeit bis zum ersten Auftreten einer hepatischen Enzephalopathie mit CTCAE Grad 3 oder 4 ergab sich ein statistisch signifikanter Nachteil für Lenvatinib gegenüber Sorafenib (HR=3,01; 95%-KI [1,29; 7,02]). Bezüglich der Zeit bis zum ersten Auftreten eines palmar-plantaren Erythrodysästhesiesyndroms mit CTCAE Grad 3 oder 4 ergab sich bei der Analyse eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,23; 95%-KI [0,13; 0,42]). Alle betrachteten Einzel-UE sind im Rahmen der Behandlung des HCC vorhersehbar und beherrschbar. Die medizinischen Maßnahmen, die beim Auftreten der jeweiligen Einzel-UE angezeigt sind, werden in der Fachinformation ausführlich beschrieben (Eisai 2018a)

Zusammenfassend zeigte Lenvatinib gegenüber Sorafenib hinsichtlich der Kategorien SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 weder einen geringeren noch größeren Schaden. Bezüglich der Kategorie UE zeigte sich eine signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Bei den ergänzenden Analysen der UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, ergab sich insgesamt weder ein geringerer noch ein größerer Schaden für Lenvatinib gegenüber Sorafenib.

Subgruppenanalyse

Bei den rein explorativen Subgruppenanalysen ist zu berücksichtigen, dass ein erhöhtes Auftreten von falsch positiven Ergebnissen aufgrund der Vielzahl der Analysen und des nicht für multiples Testen adjustierten Signifikanzniveaus zu erwarten ist.

Für den Endpunkt OS wurde kein Beleg für eine Interaktion festgestellt. Bezüglich PFS zeigte sich ein konsistentes Bild über alle Subgruppen hinweg, und somit bestätigten die Subgruppenanalysen die Überlegenheit von Lenvatinib gegenüber Sorafenib. Für die Fragebögen EORTC QLQ-C30 und -HCC18 (Morbidität) zeigte sich ebenfalls ein konsistentes Bild über alle Subgruppen. Die gefundenen Interaktionen sind ausschließlich quantitativ mit gleichgerichteten Effekten oder für alle Subgruppen nicht signifikant, so dass ein zufälliges Signal nicht ausgeschlossen werden kann. Aus diesem Grund werden zur Ableitung des Zusatznutzens hinsichtlich der Wirksamkeitsendpunkte OS und PFS und der Fragebögen EORTC QLQ-C30 und -HCC18 die Ergebnisse der Gesamtpopulation herangezogen.

Auch für die Fragebögen EORTC QLQ-C30 und -HCC18 (gesundheitsbezogene Lebensqualität) zeigte sich ein konsistentes Bild über alle Subgruppen. Die gefundenen Interaktionen sind ebenfalls ausschließlich quantitativ mit gleichgerichteten Effekten oder für alle Subgruppen nicht signifikant, so dass von einem zufälligen Signal ausgegangen werden kann. Zur Ableitung des Zusatznutzens hinsichtlich der Endpunkte zur gesundheitsbezogenen Lebensqualität EORTC QLQ-C30 und -HCC18 (gesundheitsbezogene Lebensqualität) werden deshalb die Ergebnisse der Gesamtpopulation herangezogen.

Für die Endpunkte zu Sicherheit und Verträglichkeit finden sich ebenfalls vereinzelt Belege für Interaktionen zwischen der Behandlung und den subgruppenbildenden Merkmalen. Auch hier sind die gefundenen Interaktionen ausschließlich quantitativ mit gleichgerichteten Effekten oder für alle Subgruppen nicht signifikant, so dass von einem zufälligen Signal ausgegangen werden kann. Zur Ableitung des Zusatznutzens hinsichtlich der Sicherheit und Verträglichkeit werden aus diesem Grund die Ergebnisse der Gesamtpopulation herangezogen.

Zusammenfassend konnte hinsichtlich OS, der Morbiditätsendpunkte, der Endpunkte zur gesundheitsbezogenen Lebensqualität und der Endpunkte zu Sicherheit und Verträglichkeit keine hinreichende Evidenz für eine Effektmodifikation in der Studie 304 identifiziert werden. Daher basiert die Ableitung des Zusatznutzens von Lenvatinib im Vergleich zu Sorafenib für alle berücksichtigten Endpunkte auf den Ergebnissen der Gesamtpopulation.

Die nachfolgende Tabelle zeigt die Ableitung der Aussagesicherheit und des Ausmaßes des Zusatznutzens gemäß der AM-NutzenV festgelegten Kategorien.

Tabelle 4-B: Ausmaß des Zusatznutzens von Lenvatinib vs. Sorafenib auf Endpunktebene

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
Mortalität		
OS (Hauptanalyse)	HR=0,92 [0,79; 1,06]; p=0,2526 13,6 vs. 12,3 Monate	**Hinweis / Zusatznutzen mit nicht-quantifizierbarem Ausmaß**
OS (multivariate adjustierte Analyse)	HR=0,82 [0,70; 0,96]	
OS (AFP adjustierte Analyse)	HR=0,85 [0,73; 0,99]	
Morbidität		
PFS (IIR per mRECIST)	HR=0,64 [0,55; 0,75]; p<0,0001 7,3 vs. 3,6 Monate	**Hinweis / Zusatznutzen mit Ausmaß beträchtlich**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Fatigue	HR=0,93 [0,797; 1,085]; p=0,3592 1,8 vs. 1,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Übelkeit und Erbrechen	HR=1,04 [0,857; 1,263]; p=0,6894 6,4 vs. 7,3 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Schmerz	HR=0,81 [0,690; 0,949]; p=0,0091 1,9 vs. 1,8 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Dyspnoe	HR=0,97 [0,797; 1,170]; p=0,7217 6,4 vs. 5,6 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Schlaflosigkeit	HR=1,16 [0,958; 1,398]; p=0,1293 4,6 vs. 6,1 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Appetitminderung	HR=1,02 [0,858; 1,201]; p=0,8575 2,8 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Obstipation	HR=1,06 [0,868; 1,301]; p=0,5537 7,3 vs. 9,2 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Diarrhö	HR=0,52 [0,440; 0,623]; p<0,0001 4,6 vs. 2,7 Monate	**Hinweis / Zusatznutzen mit Ausmaß erheblich**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Fatigue	HR=0,94 [0,802; 1,097]; p=0,4216 1,8 vs. 1,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Ikterus	HR=0,95 [0,790; 1,144]; p=0,5934 4,6 vs. 3,7 Monate	Kein Hinweis auf einen Zusatznutzen

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert** **mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten** **Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Schmerz	HR=1,14 [0,961; 1,349]; p=0,1315 2,8 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Fieber	HR=0,90 [0,752; 1,083]; p=0,2705 5,5 vs. 3,7 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Abdominale Schwellung	HR=0,98 [0,805; 1,197]; p=0,8549 7,4 vs. 7,4 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 7 mm im EQ-5D VAS	HR=0,90 [0,765; 1,056]; p=0,1951 2,8 vs. 1,9 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 mm im EQ-5D VAS	HR=0,89 [0,758; 1,050]; p=0,1713 2,8 vs. 1,9 Monate	Kein Hinweis auf einen Zusatznutzen
Gesundheitsbezogene Lebensqualität		
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Physische Funktion	HR=0,89 [0,756; 1,056]; p=0,1870 3,7 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Rollenfunktionsniveau	HR=0,82 [0,696; 0,962]; p=0,0148 2,0 vs. 1,9 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Emotionale Funktion	HR=1,01 [0,827; 1,235]; p=0,9142 7,4 vs. 6,6 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Kognitive Funktion	HR=1,09 [0,920; 1,291]; p=0,3167 2,8 vs. 3,0 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Soziale Funktion	HR=1,06 [0,899; 1,260]; p=0,4658 2,8 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Finanzielle Schwierigkeiten	HR=0,91 [0,735; 1,130]; p=0,3980 12,3 vs. 8,5 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im	HR=0,99 [0,841; 1,171]; p=0,9247 2,7 vs. 2,1 Monate	Kein Hinweis auf einen Zusatznutzen

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
EORTC QLQ-C30 Globaler Gesundheitsstatus		
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Körperbild	HR=0,79 [0,671; 0,932]; p=0,0051 2,8 vs. 1,9 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Ernährung	HR=0,80 [0,674; 0,949]; p=0,0102 4,3 vs. 2,8 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Sexualleben	HR=0,94 [0,771; 1,154]; p=0,5700 7,4 vs. 6,7 Monate	Kein Hinweis auf einen Zusatznutzen
Sicherheit und Verträglichkeit		
Zeit bis zum ersten UE	HR=0,84 [0,74; 0,95] 8 vs. 8 Tage	Kein geringerer oder größerer Schaden
Zeit bis zum ersten SUE	HR=1,24 [0,999; 1,53)] 412 vs. 708 Tage	Kein geringerer oder größerer Schaden
Zeit bis zum ersten UE, das zum Therapieabbruch führte	HR=1,10 [0,80; 1,50] NE vs. NE Tage	Kein geringerer oder größerer Schaden
Zeit bis zum ersten UE mit CTCAE Grad ≥ 3	HR=0,96 [0,83; 1,12] 82 vs. 64 Tage	Kein geringerer oder größerer Schaden

CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; NE: not estimable (nicht berechenbar); OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); SUE: Schwerwiegendes unerwünschtes Ereignis / schwerwiegende unerwünschte Ereignisse; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Tabelle 4-C: Übersicht zur Gesamtaussage zum Zusatznutzen von Lenvatinib vs. Sorafenib

Zusatznutzen	**Größerer Schaden**
erheblich • EORTC QLQ-C30 Symptome (Hinweis) • Diarrhö	erheblich –
beträchtlich • PFS (Hinweis)	beträchtlich –

gering
• EORTC QLQ-C30 Symptome (Hinweis)
  • Schmerz
• EORTC QLQ-C30 QoL (Hinweis)
  • Rollenfunktionsniveau
• EORTC QLQ-HCC18 QoL (Hinweis)
  • Körperbild
  • Ernährung
	gering
	–

nicht-quantifizierbar • OS	nicht-quantifizierbar –
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); OS: Overall Survival (Gesamtüberleben); PFS: Progression-free Survival (progressionsfreies Überleben); QoL: Quality of Life	

Zusammenfassung

Unter Berücksichtigung der für die direkt und unmittelbar patientenrelevanten Wirksamkeitsendpunkte OS und PFS, für die Skalen zur Morbidität des EORTC QLQ-C30 und für die Skalen zur gesundheitsbezogenen Lebensqualität des EORTC QLQ-30 und QLQ-HCC18 gezeigten Verbesserungen des therapierelevanten Nutzens für diese seltene, sehr schwere und praktisch immer unheilbare Tumorerkrankung, für die ein dringender Therapiebedarf bei gleichzeitig sehr stark beschränkter Anzahl an Therapieoptionen besteht, ergibt sich in der Zusammenschau mit dem insgesamt vergleichbaren, vorhersehbaren und beherrschbaren Sicherheits- und Verträglichkeitsprofil für Lenvatinib gegenüber der ZVT Sorafenib ein **Hinweis auf einen erheblichen Zusatznutzen**.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

– Patientenpopulation
– Intervention
– Vergleichstherapie
– Endpunkte
– Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Fragestellung

Lenvatinib ist zugelassen als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem hepatozellulärem Karzinom (HCC), die zuvor noch keine systemische Therapie erhalten haben (Eisai 2018a).

Das final zugelassene Anwendungsgebiet unterscheidet sich von dem zum Zeitpunkt der Beratungsanforderung bei der European Medicines Agency (EMA) eingereichten Anwendungsgebiet, das wie folgt lautete: Lenvatinib ist zugelassen für erwachsene Patienten mit Leberzellkarzinom. Für dieses ursprüngliche Anwendungsgebiet hat der Gemeinsame Bundesausschuss (G-BA) folgende zweckmäßige Vergleichstherapien (ZVT) bestimmt (G-BA 2018a):

Die zweckmäßige Vergleichstherapie für Lenvatinib zur Behandlung des Leberzellkarzinoms für Patienten ohne kurative Therapieintention und für die eine lokoregionäre Therapie nicht in Frage kommt, ist:

a) für Patienten mit Barcelona Clinic Liver Cancer (BCLC)-Stadium B oder C

   a1) mit Child-Pugh A oder keine Leberzirrhose: Sorafenib

a2) mit Child-Pugh B oder mit Sorafenib vorbehandelt: Best-Supportive-Care

b) für Patienten mit BCLC-Stadium D: Best-Supportive-Care

Nach der Beratungsanforderung wurde das Anwendungsgebiet von Lenvatinib im Laufe des Zulassungsverfahrens folgendermaßen spezifiziert: Lenvatinib ist indiziert als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben (Eisai 2018a). Wie in Modul 3.1. erläutert, ergibt sich aus der Spezifizierung, dass die vom G-BA definierte Population b (Patienten im BCLC-Stadium D) sowie die Patienten in Population a2, die mit Sorafenib vorbehandelt sind, außerhalb des Anwendungsgebietes liegen.

Studie 304 ist geeignet, den Nachweis der Wirksamkeit und Verträglichkeit von Lenvatinib gegenüber Sorafenib in den Populationen a1 und a2 zu erbringen (siehe Modul 3.1). Die Patienten befanden sich bei Einschluss in die Studie ausnahmslos in den BCLC Stadien B und C und wiesen einen Child-Pugh Score von 5 bis 8 auf, dies entspricht den Einteilungen Child-Pugh A und Child-Pugh B. Dabei muss berücksichtigt werden, dass der Großteil der Patienten einen Child-Pugh Score von 5 oder 6 (entspricht Child-Pugh A) aufwiesen.

Der G-BA hat für Patienten, die einen Child-Pugh Score von 7 bis 9 aufwiesen, Best Supportive Care (BSC) als ZVT bestimmt. Dieser Einteilung kann sich Eisai insofern anschließen, dass für Patienten mit einem Child-Pugh Score von 7 bis 9, für die keine systemische Therapie mehr infrage kommt, BSC eine angemessene ZVT ist. Da diese Patienten nicht den Allgemeinzustand aufweisen, der eine systemische Therapie ermöglicht, ist auch die Teilnahme an der Zulassungsstudie für ein neues Medikament mit einer bis zu diesem Zeitpunkt nicht nachgewiesenen Wirksamkeit nicht vertretbar. Zusätzlich gibt es auch Patienten mit einem Child-Pugh Score von 7 bis 9, die noch für eine systemische Therapie in Frage kommen. In diesem Fall muss eine systemische Therapie die ZVT sein. Etwas anderes wäre im Rahmen von Zulassungsstudien ethisch nicht vertretbar. Dementsprechend sollte die Studie 304 zum Nachweis des Zusatznutzens von Lenvatinib gegenüber Sorafenib für all jene Patienten herangezogen werden, die weiterhin für einen systemischen Therapieansatz in Frage kommen und nicht bereits mit Sorafenib vorbehandelt sind, unabhängig vom Child-Pugh Score des einzelnen Patienten.

Daraus ergibt sich die folgende Fragestellung:

Fragestellung: Wie ist die quantitative und qualitative Evidenz aus RCT (Randomized Controlled Trial(s), randomisierte(n) kontrollierte(n) Studie(n)) hinsichtlich eines Zusatznutzens von Lenvatinib als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben, im Vergleich zu Sorafenib hinsichtlich der patientenrelevanten Endpunkte Mortalität (Gesamtüberleben), Morbidität und gesundheitsbezogene Lebensqualität sowie der patientenrelevanten Endpunkte zu Sicherheit und Verträglichkeit zu bewerten?

Patientenpopulationen

Lenvatinib (Lenvima®) ist laut Fachinformation (Eisai 2018a) zugelassen als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben. Dabei muss berücksichtigt werden, dass Lenvatinib, wie in Modul 3 ausführlich dargestellt, als systemische Therapie in der Versorgungsrealität nur bei Patienten eingesetzt wird, die für eine systemische Therapie in Frage kommen.

Intervention

Die für die Nutzenbewertung relevante Therapie ist Lenvatinib, ein Rezeptor-Tyrosinkinase (RTK)-Inhibitor, der selektiv die Kinaseaktivitäten der VEGF (engl. Vascular Endothelial Growth Factor, vaskulären endothelialen Wachstumsfaktor)- Rezeptoren VEGFR1 , VEGFR2 und VEGFR3 sowie andere, mit dem proangiogenen und onkogenen Signalweg in Zusammenhang stehende RTKs, einschließlich der Rezeptoren FGFR1, 2, 3 und 4 des Fibroblasten-Wachstumsfaktors (engl. Fibroblast Growth Factor, FGF) und die Rezeptoren PDGFRα, KIT und RET des Blutplättchen-Wachstumsfaktors (engl. Platelet-Derived Growth Factor, PDGF), hemmt. Lenvatinib ist der erste Wirkstoff, der gleichzeitig sowohl die Tätigkeit der RTK des VEGF als auch des FGF hemmt. Der Wirkmechanismus von Lenvatinib ist in Modul 2 dieses Nutzendossiers ausführlich dargestellt.

Vergleichstherapie

Die Bewertung des Zusatznutzens von Lenvatinib als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben, erfolgt gegenüber der vom G-BA bestimmten ZVT Sorafenib als systemische Therapie.

Der Nachweis des Nutzens und des Zusatznutzens von Lenvatinib erfolgt im Rahmen eines direkten Vergleichs von Lenvatinib gegenüber der systemischen Therapie mit Sorafenib basierend auf den Ergebnissen der randomisierten kontrollierten pivotalen Phase-III-Studie 304.

Endpunkte und Studientypen

Mittels geeigneter, patientenrelevanter Endpunkte, die im Rahmen eines methodisch hochwertigen RCT betrachtet werden, wird der Nutzen sowie der Zusatznutzen hinsichtlich patientenrelevanter therapeutischer Endpunkte (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verbesserung der Lebensqualität oder Verringerung von Nebenwirkungen) untersucht.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Population

Die Studienpopulation umfasst erwachsene Patienten mit fortgeschrittenem oder inoperablenHCC, die zuvor noch keine systemische Therapie erhalten haben (Eisai Inc. 2017b). Zur Definition des fortgeschrittenen oder inoperablen Leberzellkarzinoms siehe Modul 3.2.1.

Intervention

Die Intervention ist Lenvatinib (Lenvima®), ein oraler RTK-Inhibitor, der gemäß Fachinformation als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben, angewendet wird. Die Bewertung erfolgt in der zugelassenen Dosierung Lenvatinib 12 mg (drei 4 mg-Kapseln einmal täglich) für Patienten mit einem Körpergewicht ≥ 60 kg zu Baseline und Lenvatinib 8 mg für Patienten mit einem Körpergewicht < 60 kg zu Baseline (zwei 4 mg-Kapseln einmal täglich) (Eisai 2018a).

Vergleichstherapie

Die relevante Vergleichstherapie ist Sorafenib (Nexavar®) in der gemäß Fachinformation zugelassenen Dosierung (zwei 200 mg Tabletten zweimal täglich) (Bayer Pharma AG 2018).

Endpunkte

Die folgenden Endpunkte werden zur Beurteilung patientenrelevanter therapeutischer Effekte herangezogen:

Mortalität

- Gesamtüberleben (engl. Overall Survival, OS)

Morbidität

- Progressionsfreies Überleben (engl. Progression Free Survival, PFS)

- Fragebogen zur Lebensqualität der Europäische Organisation für Krebsforschung und -behandlung (engl. European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire, EORTC QLQ-C30) bezüglich der Symptome:
    - Fatigue
    - Übelkeit und Erbrechen
    - Schmerz
    - Dyspnoe
    - Schlaflosigkeit
    - Appetitminderung
    - Obstipation
    - Diarrhö
- HCC-spezifischer Fragebogen zur Lebensqualität der Europäischen Organisation für Krebsforschung und -behandlung (engl. European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire, EORTC QLQ-HCC18) bezüglich der Symptome:
    - Fatigue
    - Ikterus
    - Schmerz
    - Fieber
    - Abdominale Schwellung
- Fragebogen zur Lebensqualität der EuroQoL Gruppe auf einer visuellen Analogskala (engl. EuroQoL Group - Quality of Life Questionnaire on a vertical visual analogue scale, EQ-5D VAS)

Gesundheitsbezogene Lebensqualität

- EORTC QLQ-C30 bezüglich der Funktionen:
    - Physische Funktion
    - Rollenfunktionsniveau
    - Emotionale Funktion
    - Kognitive Funktion
    - Soziale Funktion
    - Globaler Gesundheitsstatus
    - Finanzielle Schwierigkeiten
- EORTC QLQ-HCC18 bezüglich der Items:

- Körperbild
- Ernährung
- Sexualleben

Sicherheit und Verträglichkeit

- Unerwünschte/s Ereignisse (UE)
- Schwerwiegende/s UE (SUE)
- UE, die zum Therapieabbruch führten
- UE mit Common Terminology Criteria for Adverse Events (CTCAE) Grad ≥ 3

Studientypen

Die Fragestellung zur Bewertung des medizinischen Zusatznutzens von Lenvatinib kann mittels prospektiver RCT am besten beantwortet werden. In der Selektion gilt dieser Studientyp als Einschlusskriterium.

Studiendauer

Es werden keine Einschränkungen bezüglich der Behandlungsdauer gemacht, da die Dauer ein Teil der Zielgröße OS darstellt.

Tabelle 4-D: Übersicht der Ein- und Ausschlusskriterien der Studien, die in die Nutzenbewertung eingehen

Kategorie	**Einschlusskriterien**		**Ausschlusskriterien**	
Population	E1	▪ Erwachsene Patienten mit fortgeschrittenem oder inoperablem hepatozellulärem Karzinom (HCC), die zuvor noch keine systemische Therapie erhalten haben	A1	▪ Patienten < 18 Jahre, die nicht E1 entsprechen
Intervention	E2	▪ Patienten mit einem Körpergewicht ≥ 60 kg zu Baseline: Lenvatinib (12 mg), oral ▪ Patienten mit einem Körpergewicht < 60 kg zu Baseline: Lenvatinib (8 mg), oral	A2	▪ Lenvatinib in anderer als der gemäß Fachinformation empfohlenen Tagesdosis von Lenvatinib (12 mg für Patienten mit einem Körpergewicht ≥ 60 kg oder 8 mg für Patienten mit einem Körpergewicht < 60 kg)
Vergleichstherapie	E3	▪ Sorafenib in zugelassener Tagesdosis (800 mg), oral	A3	▪ Sorafenib in nicht zugelassener Tagesdosis
Patientenrelevante Endpunkte	E4	Mindestens einer der gelisteten Endpunkte: ▪ OS ▪ PFS ▪ Morbidität (Symptome) ▪ Gesundheitsbezogene Lebensqualität ▪ Unerwünschte Ereignisse	A4	▪ Andere Endpunkte als in E4 ▪ Punkt- und Streumaße nicht vorhanden
Studientypen	E5	▪ RCT	A5	▪ Tierexperimentelle Studien ▪ Keine RCT
Studiendauer	E6	▪ Keine Einschränkung bezüglich der Studiendauer	A6	▪ Keine Einschränkung bezüglich der Studiendauer
Sprache	E7	▪ Englisch oder Deutsch	A7	▪ Andere als Englisch oder Deutsch
Publikationstyp	E8	▪ Vollpublikation zu RCT	A8	▪ Mehrfachpublikationen ohne relevante Zusatzinformationen ▪ Kein Studienbericht
Studienstatus	E9	▪ Studien, für die Ergebnisse vorliegen	A9	▪ Noch nicht begonnene, laufende oder abgeschlossene Studien, für die noch keine Ergebnisse vorliegen

kg: Kilogramm; mg: Milligramm; OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 0, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.3.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.3.2) sowie weitere Untersuchungen (Abschnitt 4.3.3.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die bibliografische Literaturrecherche erfolgte am 13. Juli 2018. Die Recherche wurde in den Datenbanken

- MEDLINE
- EMBASE
- Cochrane Central Register of Controlled Trials
- Cochrane Database of Systematic Reviews

über die Plattform Ovid durchgeführt. Für die Eingrenzung auf den Studientyp 'Randomisierte kontrollierte Studien' wurde für EMBASE und MEDLINE der validierte Filter nach Wong verwendet (Wong et al. 2006).

Die Suchstrategie setzte sich aus einer Kombination von Suchwortblöcken zu Indikation, Studiendesign sowie Intervention zusammen und wurde an die jeweilige Datenbank angepasst. Zeitliche Einschränkungen wurden nicht vorgenommen. Berücksichtigt wurden Publikationen in deutscher und englischer Sprache.

Detaillierte Angaben zu den Suchstrategien, dem Rechercheverlauf und den verwendeten Suchbegriffen finden sich in Anhang 4-A. Ergebnisse zu den Recherchen finden sich in Abschnitt 4.3.1.1.2 und Anhang 4-C.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.3.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.3.2) sowie weitere Untersuchungen (Abschnitt 4.3.3.3) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: http://apps.who.int/trialsearch/) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in 0 zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Die Suche nach Studien mit dem zu bewertenden Arzneimittel Lenvatinib wurde am 06. Juli 2018 in den öffentlich zugänglichen Studienregistern

- clinicaltrials.gov (www.clinicaltrials.gov)
- EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu)
- Klinische Prüfungen PharmNet.Bund (http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm)
- International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: http://apps.who.int/trialsearch/)

durchgeführt. Die Suche wurde an das jeweilige Studienregister angepasst. Zusätzliche Einschränkungen der Suche wurden nicht vorgenommen. Eine detaillierte Beschreibung der Studienregisterrecherche findet sich in 0. Die Ergebnisse der Suche in den Studienregistern finden sich in dem Abschnitt 4.3.1.1.3.

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechercheschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Selektion der Publikationen und Studien aus den Ergebnissen der bibliografischen Literaturrecherche und der Suche in Studienregistern erfolgte anhand der im Abschnitt 4.2.2 beschriebenen Ein- und Ausschlusskriterien. Die Prüfung der erhaltenen Treffer aus der bibliografischen Recherche auf ihre Relevanz erfolgte in einem ersten Schritt auf Basis des Titels und des Abstracts, anschließend wurde bei als potenziell relevant erachteten Publikationen der Volltext gesichtet. Die aus den Studienregisterrecherchen identifizierten Studien wurden anhand der Informationen aus den Studienregistern hinsichtlich ihrer Relevanz beurteilt. Die Bewertung über den Ein- oder Ausschluss wurde von zwei Gutachtern unabhängig voneinander vorgenommen. Bei voneinander abweichenden Einschätzungen beider Gutachter wurde durch Diskussion eine einheitliche Bewertung erzielt.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

– Erzeugung der Randomisierungssequenz *(bei randomisierten Studien)*
– Verdeckung der Gruppenzuteilung *(bei randomisierten Studien)*
– zeitliche Parallelität der Gruppen *(bei nicht randomisierten vergleichenden Studien)*
– Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren *(bei nicht randomisierten vergleichenden Studien)*
– Verblindung des Patienten sowie der behandelnden Personen
– ergebnisgesteuerte Berichterstattung
– sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

– Verblindung der Endpunkterheber
– Umsetzung des ITT-Prinzips
– ergebnisgesteuerte Berichterstattung
– sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Bewertung erfolgte auf Studienebene und separat für jeden patientenrelevanten Endpunkt auf Basis des klinischen Studienberichts. Die Extraktion und Dokumentation wurde unter Verwendung von Anhang 4-F und mittels der folgenden endpunktübergreifenden und endpunktspezifischen Aspekte durchgeführt:

- Verzerrungsaspekte auf Studienebene (A)
  - Erzeugung der Randomisierungssequenz
  - Verdeckung der Gruppenzuteilung
  - Verblindung der Patienten sowie der Behandler
  - Ergebnisgesteuerte Berichterstattung
  - Sonstige Verzerrungsaspekte
- Verzerrungsaspekte der Ergebnisse auf Endpunktebene (B)
  - Verblindung der Endpunkterheber
  - Umsetzung des Intention-to-treat (ITT)-Prinzips
  - Ergebnisgesteuerte Berichterstattung
  - Sonstige Verzerrungsaspekte

Die Beurteilung des Verzerrungspotenzials, das aus der Selektion der Studienpopulation resultiert, wurde anhand der Randomisierung der Patienten vorgenommen. Dabei mussten die Erzeugung der Randomisierungssequenz und die daraus resultierende Verteilung der Patienten adäquat durchgeführt und dokumentiert worden sein. Ebenfalls musste die Randomisierungssequenz den Behandlern verborgen geblieben sein. Um eine Verzerrung der Studienergebnisse durch den Patienten oder den Behandler zu vermeiden, sollten beide in adäquater Weise verblindet worden sein. Bei offen durchgeführten Studien ist die mögliche

Auswirkung auf das Verzerrungspotenzial zu diskutieren. Des Weiteren wurde kritisch überprüft, ob die gültigen Kriterien für das Studiendesign eingehalten wurden.

Um eine Verzerrung auf Ebene der Studienendpunkte ausschließen zu können, mussten alle festgelegten Zielgrößen vollständig berichtet werden. Dabei wird die ITT-Population als die Studienpopulation mit dem geringsten Verzerrungspotenzial betrachtet. Die Einhaltung des ITT-Prinzips wurde im Einklang mit den Vorgaben des vorliegenden Dokuments u.a. dann als adäquat umgesetzt bewertet, wenn der Nichtberücksichtigungsanteil (von z.B. Protokollverletzern und Lost-to-Follow-up-Patienten) in der Auswertung < 5% betrug.

Eine zusammenfassende Bewertung des Verzerrungspotenzials auf Studien- und Endpunktebene erfolgte in den Kategorien der Ergebnissicherheit „hoch“ und „niedrig“. Ein niedriges Verzerrungspotenzial lag dann vor, wenn davon ausgegangen werden konnte, dass die Ergebnisse nicht relevant verzerrt waren. Eine relevante Verzerrung ist dann anzunehmen, wenn sich die Ergebnisse insgesamt oder die Ergebnisse auf Endpunktebene grundlegend verändern.

Wurde das Verzerrungspotenzial auf Studienebene (A) mit „hoch“ bewertet, erfolgte die Einstufung des Verzerrungspotenzials auf Endpunktebene (B) ebenfalls als „hoch“. Ein hohes Verzerrungspotenzial führte nicht zum Ausschluss der Daten.

Das Ausmaß der qualitativen Ergebnissicherheit auf Studien- und Endpunktebene wurde anschließend auf Basis des Verzerrungspotenzials und der European Medicines Agency (Europäische Arzneimittel Agentur, EMA)-Kriterien für die Zulassung auf Grundlage einer einzelnen pivotalen Studie bestimmt (EMA 2001). Die letztgenannten Kriterien werden durch die Erteilung einer Zulassung für Lenvima® zur Behandlung des HCC auf Basis einer Phase-III-Studie als gegeben betrachtet.

Im Methodenpapier 5.0 des Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) werden die Kategorien zur Graduierung des Ausmaßes beschrieben (IQWiG 2017):

- **Hohe Ergebnissicherheit:** Ergebnis bezüglich eines Endpunkts einer randomisierten Studie mit niedrigem Verzerrungspotenzial.
- **Mäßige Ergebnissicherheit:** Ergebnis bezüglich eines Endpunkts einer randomisierten Studie mit hohem Verzerrungspotenzial.
- **Geringe Ergebnissicherheit:** Ergebnis bezüglich eines Endpunkts einer nicht randomisiert vergleichenden Studie.

Die weitere Beurteilung umfasste die Bewertung der Aussagesicherheit auf Grundlage der Anzahl vorhandener Studien, deren Ergebnissicherheit und der gemessenen Effekte in den Kategorien Beleg, Hinweis, Anhaltspunkt oder kein nachgewiesener Zusatznutzen (IQWiG 2017).

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 0, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)[1]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[2] bzw. STROBE-Statements[3] folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Studiendesign und Methodik der eingeschlossenen RCT wurde anhand der Vorgaben des CONSORT (Consolidated Standards of Reporting Trials)-Statements (Items 2b bis 14 und Flow-Chart) beschrieben (siehe Anhang 4-E).

Das CONSORT-Statement diente als Checkliste (Schulz et al. 2010) und erforderte Ausführungen zu jedem der gebotenen Items (Moher et al. 2010). Das Ziel der Studie, die dazugehörigen Hypothesen, das Studiendesign, die Studienpopulation, die Interventionen und die betrachteten Zielgrößen wurden in Item 2b bis 6 beschrieben. Die Fallzahlplanung, das Vorgehen zur Randomisierung, die Generierung der Randomisierungssequenz, die verdeckte Zuteilung der Patienten zu den Studienarmen, die Verblindung und die adäquate Anwendung statistischer Analysemethoden wurden in Item 7 bis 12 dargestellt. Die Bewertung dargestellter Ergebnisse wurde anhand der Zahl der eingeschlossenen, randomisierten, behandelten und analysierten Patienten beurteilt. Studienabbrüche und Ausschlüsse von Patienten nach der Randomisierung wurden präzise beschrieben. Der Durchführungszeitraum der Studie, die Nachbeobachtung der Patienten und die Beendigung der Studie wurden ebenso exakt in Item 13 und 14 beschrieben.

[1] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[2] Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.
[3] Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 0 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Patientencharakteristika

Die Studienpopulation umfasst erwachsene Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben. Die nachfolgende Tabelle gibt einen Überblick über die Patientencharakteristika der Patienten in der Studie 304 zu Baseline.

Tabelle 4-E: Patientencharakteristika der Studie 304 zu Baseline

Kategorie	Ausprägung
Baseline Alter (kontinuierlich)	▪ Jahre
Baseline Altersgruppe (kategorial)	▪ < 65 Jahre ▪ ≥ 65 bis < 75 Jahre ▪ ≥ 75 Jahre
Geschlecht (kategorial)	▪ Männlich ▪ Weiblich
Region[a] (kategorial)	▪ Asien-Pazifik ▪ Westliche Regionen
Ethnische Gruppe (kategorial)	▪ Weiße ▪ Schwarze oder Afroamerikaner ▪ Asiaten o Anteil Japaner o Anteil Chinesen o Anteil andere Asiaten ▪ Indianer oder alaskanische Ureinwohner

Kategorie	Ausprägung
	▪ Eingeborene Hawaiianer oder andere pazifische Inselbewohner ▪ Andere
Ethnizität (kategorial)	▪ Spanisch oder lateinamerikanisch ▪ Nicht spanisch oder lateinamerikanisch
Makroskopische Invasion der Pfortader (MPVI) (kategorial)	▪ Ja ▪ Nein
Extrahepatische Ausbreitung (ES) (kategorial)	▪ Ja ▪ Nein
MPVI oder ES oder beides (kategorial)	▪ Ja ▪ Nein
Baseline ECOG-PS (kategorial)	▪ 0 ▪ 1
Baseline Körpergewicht (kontinuierlich)	▪ kg
Baseline Gewichtsgruppe (kategorial)	▪ < 60 kg ▪ ≥ 60 kg
Baseline Körpergröße (kontinuierlich)	▪ cm
NYHA-Klassifikation (kategorial)	▪ I ▪ II ▪ Nicht zutreffend ▪ Unbekannt
Child-Pugh Score (kategorial)	▪ 5 ▪ 6 ▪ 7 ▪ 8
Zugrundeliegende Leberzirrhose (kategorial)	▪ Ja ▪ Nein
Zeit seit erster Diagnose (kontinuierlich)	▪ Monate
Alter bei erster Diagnose (kontinuierlich)	▪ Jahre
BCLC Staging (kategorial)	▪ Stadium B ▪ Stadium C
Betroffene Krankheitsstellen (kategorial)	▪ Leber ▪ Lunge ▪ Lymphknoten ▪ Knochen ▪ Andere
Anzahl der betroffenen Krankheitsstellen pro Patient (kategorial)	▪ 1 ▪ 2

Kategorie	Ausprägung
	▪ ≥ 3
Faktor der Karzinogenese (kategorial)	▪ HBV ▪ HCV ▪ Alkohol ▪ Andere ▪ Unbekannt
Leberzirrhose zu Baseline (kategorial)	▪ Ja ▪ Nein
Baseline AFP (kontinuierlich)	▪ ng/mL
Baseline AFP (kategorial)	▪ < 200 ng/mL ▪ ≥ 200 ng/mL
Ammoniakspiegel (kontinuierlich)	▪ µg/dL
Begleitende systemische antivirale Therapie von HBV oder HCV (kategorial)	▪ Frühere Therapien zur Krebsbekämpfung ▪ Strahlentherapie
a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die Asien-Pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; cm: Zentimeter; dL: Deziliter; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; kg: Kilogramm; µg: Mikrogramm; mL: Milliliter; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); ng: Nanogramm; NYHA: New York Heart Association	

Patientenrelevante Endpunkte

Als primärer Endpunkt wurde in der Studie 304 das OS untersucht. Als sekundäre Endpunkte wurden unter Anderem die Daten für das PFS, für die EORTC QLQ-C30 Symptom- und Funktionsskalen, für die Skalen des EORTC QLQ-HCC18 und für die EQ-5D VAS erhoben. Des Weiteren wurden Sicherheit und Verträglichkeit der Behandlung als patientenrelevante Endpunkte untersucht.

Im vorliegenden Nutzendossier werden die folgenden patientenrelevanten Endpunkte berichtet:

Mortalität

- OS

Morbidität

- PFS
- EORTC QLQ-C30 (Symptome)

    - Fatigue
    - Übelkeit und Erbrechen
    - Schmerz
    - Dyspnoe
    - Schlaflosigkeit
    - Appetitminderung
    - Obstipation
    - Diarrhö
- EORTC QLQ-HCC18 (Symptome)
    - Fatigue
    - Ikterus
    - Schmerz
    - Fieber
    - Abdominale Schwellung
- EQ-5D VAS

Gesundheitsbezogene Lebensqualität

- EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität)
    - Physische Funktion
    - Rollenfunktionsniveau
    - Emotionale Funktion
    - Kognitive Funktion
    - Soziale Funktion
    - Globaler Gesundheitsstatus
    - Finanzielle Schwierigkeiten
- EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität)
    - Körperbild
    - Ernährung
    - Sexualleben

Sicherheit und Verträglichkeit

- UE

- SUE
- UE, die zum Therapieabbruch führten
- UE mit CTCAE Grad ≥ 3

Die Objektive Ansprechrate (engl. Objective Response Rate, ORR) wird in diesem Dossier als ergänzende Analyse berichtet, aber nicht zur Ableitung des Zusatznutzens herangezogen.

Mortalität

Gesamtüberleben (Overall Survival, OS)

OS wurde definiert als die Zeitspanne zwischen dem Zeitpunkt der Randomisierung bis zum Eintreten des Todes aufgrund jeglicher Ursache. Lost-to-follow-up-Patienten wurden zum letzten Zeitpunkt, an dem sie bekanntermaßen am Leben waren, zensiert. Überlebende Patienten wurden zum Zeitpunkt des Datenschnitts zensiert (Eisai Inc. 2017b).

Gemäß des Rapid Reports des IQWiG stellt das OS ein herausragendes Therapieziel in der Behandlung onkologischer Erkrankungen dar. Entsprechend ist der Endpunkt OS als unmittelbar patientenrelevant anzusehen (IQWiG 2011).

Morbidität

Progressionsfreies Überleben (Progression-free survival, PFS)

PFS ist definiert als Zeitspanne zwischen dem Zeitpunkt der Randomisierung bis zum Zeitpunkt einer dokumentierten Krankheitsprogression oder bis zum Zeitpunkt des Todes aufgrund jeglicher Ursache, abhängig davon, welches dieser beiden Ereignisse zuerst eintrat (Eisai Inc. 2017b). Somit dokumentiert PFS den direkten Effekt einer Behandlung auf die Tumorprogression unter Verwendung objektiver und quantitativer Methodik und ist als wertvoller Endpunkt für klinische Studien mit krankheitsmodifizierend wirkenden Substanzen anerkannt (FDA 2007). Der Zeitpunkt der objektiven Krankheitsprogression war in der Studie 304 definiert als der früheste Zeitpunkt einer radiologisch bestätigten Krankheitsprogression. Patienten, die die Therapie aufgrund einer Krankheitsprogression beendeten oder bei denen keine radiologische Bestätigung für die Krankheitsprogression vorlag, wurden zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert. Patienten, die eine neue Therapie anfingen oder die nach mehr als einer verpassten Visite oder nach bis zu 28 Tagen der letzten Dosis der Studienmedikation verstarben oder bei denen eine Progression auftrat, wurden ebenfalls zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert (Eisai Inc. 2016).

Die Beurteilung der Tumorprogression erfolgte in der vorliegenden Studie 304 sowohl mittels verblindeter und unabhängiger Bewertung radiologischer Verfahren (engl. Independent Imaging Review, IIR) gemäß den speziell für das HCC entwickelten modified Response Evaluation Criteria in Solid Tumors (mRECIST)-Kriterien sowie den Response Evaluation Criteria in Solid Tumors (RECIST)-Kriterien v.1.1. Desweiteren erfolgte eine Beurteilung durch den Prüfarzt gemäß den mRECIST-Kriterien.

Bei den RECIST-Kriterien handelt es sich um eine etablierte Methodik, um über die Größenänderung der Zielläsionen die Tumorantwort auf die Therapie zu bewerten. Dabei wird die Größe der Zielläsionen mit einer einfachen linearen Summation gemessen, wobei nicht zwischen nekrotischem und lebensfähigem Gewebe unterschieden wird. In mehreren HCC-Studien wurde allerdings nur eine schwache Korrelation zwischen dem klinischen Nutzen neuer systemischer Therapien wie Sorafenib oder lokoregionärer Therapien und der konventionellen radiologischen Messung der Tumorantwort nachgewiesen (Lencioni und Llovet 2010). Limitationen der Messung der Tumorantwort mithilfe der RECIST-Kriterien ergeben sich aus (1) dem möglicherweise fehlenden Umbau von zirrhotischem Lebergewebe um den Tumor herum, (2) der diffusen, infiltrierenden Beschaffenheit vom HCC in vielen zirrhotischen Lebern, (3) der Veränderung der Tumor-Vaskularität, aber nicht der Tumorgröße und (4) der verstärkten Darstellung prämaligner dysplastischer Knoten während der arteriellen Phase und sich daraus ergebender falsch-positiver Bestimmung eines Tumorwachstums (Thomas et al. 2010). Wie bereits im Jahr 2000 von der European Association for the Study of the Liver (EASL) festgestellt wurde, spiegelt sich beim HCC eine umfangreiche Tumor-Nekrose nicht zwangsläufig in einer Reduktion des Umfangs der Läsion wieder (Bruix et al. 2001). Daher wurde von einem EASL-Expertengremium das Konzept der lebensfähigen Zielläsionen, die mittels kontrastverstärkender radiologischer Bildgebung identifiziert werden, vorgeschlagen (Bruix et al. 2001). Dieses Konzept der lebensfähigen Zielläsionen wurde auch in die mRECIST-Kriterien mit aufgenommen (Lencioni und Llovet 2010). Bei der mRECIST-Methodik wird nekrotisches Tumorgewebe durch die fehlende Kontrastbildung innerhalb des Tumors bei der Bildgebung identifiziert.

Da die Veränderungen der intrahepatischen Läsionen durch mRECIST über die Messung der nur lebensfähigen Bereiche der Läsionen besser erfasst werden als über RECIST und die damit einhergehende Messung der gesamten Läsion, wurden in der Studie 304 mRECIST-Kriterien für die Auswertung durch den Prüfarzt herangezogen. Diese Entscheidung erfolgte in Abstimmung mit dem Steering Committee und nach Zustimmung durch die regulatorischen Behörden sowie unter Betrachtung der wissenschaftlichen Evidenz zum Zeitpunkt der Studienplanung (Eisai Inc. 2017b). Für den IIR wurden sowohl die RECIST als auch die mRECIST Kriterien herangezogen.

Gemäß mRECIST ist eine Krankheitsprogression definiert als eine mindestens 20%-ige Zunahme der Summe der Durchmesser der lebensfähigen Zielläsionen, wobei die kleinste dokumentierte Summe der lebensfähigen Zielläsionen seit Behandlungsbeginn als Referenz dient. Eine Zielläsion gemäß mRECIST ist dabei definiert – wie gemäß RECIST – als eine messbare Läsion mit mindestens 1 cm Durchmesser in mindestens einer Dimension (Lencioni und Llovet 2010). Darüber hinaus sollte die Läsion für wiederholte Messungen geeignet sein und eine intratumorale arterielle Erweiterung bei kontrastverstärkender Computertomografie (CT) oder Magnetic Resonance Imaging (MRI) zeigen (Lencioni und Llovet 2010). Bei Anwendung der mRECIST Kriterien werden hepatische Läsionen gemäß mRECIST und nicht-hepatische Läsionen gemäß den RECIST v.1.1 Kriterien beurteilt. Im Gegensatz zu mRECIST geht gemäß RECIST v.1.1 die komplette Läsion in die Bildung der Summe der Durchmesser mit ein (Lencioni und Llovet 2010).

Die in der Studie 304 herangezogenen mRECIST und RECIST-Kriterien stellen einen anerkannten und etablierten Standard zur Bestimmung des Tumoransprechens beim HCC dar (Lencioni und Llovet 2010).

Kommt es zu einer Tumorprogression, zu Metastasen oder zu einem Rezidiv, ist dies für den betroffenen Patienten schwerwiegend (Berg et al. 2011). Es kommt dabei zur Verschlechterung bzw. zum Auftreten neuer krankheitsspezifischer Symptome sowie zu einer starken psychischen Belastung, z.B. der sog. Progredienzangst (Herschbach et al. 2001; Herschbach und Dinkel 2014; Mehnert et al. 2006). Hierbei handelt es sich um eine reaktive und bewusst wahrgenommene Furcht, die im Rahmen einer schweren, potenziell lebensbedrohlichen, oder zur Behinderung führenden Erkrankung und ihrer Behandlung entsteht (Hesse et al. 2016). Diese Progredienzangst und deren mögliche Behandlung spielt eine immer bedeutendere Rolle im Rahmen einer palliativen Versorgung von Patientinnen und Patienten (Hesse et al. 2016).

Eine Verlängerung des PFS führt zu einem verzögerten Eintreten der Progredienzangst und der krankheitsspezifischen Symptome, was wiederum eine Steigerung der Lebensqualität durch einen positiven Einfluss auf das physische und psychische Wohlbefinden der Patienten mit sich bringt. Das Hinauszögern einer Progression der Erkrankung wird daher von den Patienten als direkter Vorteil wahrgenommen und stellt für den Patienten einen selbst erlebbaren Therapieerfolg dar.

Nicht zuletzt deswegen wurde PFS in den letzten Jahren als Endpunkt von der EMA und der Food and Drug Administration (FDA) bei sehr unterschiedlichen onkologischen Arzneimitteln und Therapieindikationen im Rahmen klinischer Studien als patientenrelevanter Endpunkt für die Zulassung neuer Arzneimittel akzeptiert und bestätigt. Auch das Bundesinstitut für Arzneimittel (BfArM) erachtet das PFS als klinisch relevanten Endpunkt (BfArM 2011). Ein länger anhaltendes PFS wird dabei als relevanter Endpunkt von Patientennutzen anerkannt, wenn das Ausmaß des Behandlungseffektes ausreichend groß ist, um etwaige Verträglichkeitsprobleme zu überwiegen (EMA 2012). Das Committe for Medicinal Products for Human Use (CHMP) der EMA forderte demnach bereits im Jahr 2012 in einer Stellungnahme zur Überarbeitung der „Anticancer Guideline“, dass die alleinstehende klinische Relevanz von PFS anerkannt wird. Die CHMP Scientific Advisory Group (SAG) for Oncology, in der ausgewählte europäische Experten Änderungen der Anticancer Guideline berieten, stellten fest, dass PFS als klinisch relevanter Endpunkt zu betrachten ist, selbst wenn Verbesserungen im PFS nicht mit einer dokumentierten Verlängerung des OS einhergingen, solange es keine Hinweise auf negative Auswirkungen auf das Gesamtüberleben gibt. Dies wurde hauptsächlich begründet mit der Wichtigkeit, die das Hinauszögern des erstmaligen Auftretens oder die Verschlechterung von Symptomen in Verbindung mit einer Progression sowie einer nächsten Therapielinie für den Patienten besitzt (EMA 2012). Somit ist PFS explizit nicht als Surrogat für OS, sondern als alleinstehender, patientenrelevanter Endpunkt zu betrachten.

Auch diverse Fachgesellschaften teilen diese Meinung. Die Arbeitsgemeinschaft Dermatologische Onkologie (ADO) definiert ein länger anhaltendes mehrmonatiges PFS und den damit einhergehenden Abstand zu einer lebensbedrohlichen Situation als

patientenrelevanten Endpunkt. Eine Zeit ohne Progression und damit eine Zeit ohne Todesängste sieht die ADO eindeutig als relevantes Ziel für den Patienten (ADO 2012). Nach Einschätzung der Deutschen Gesellschaft für Hämatologie und Onkologie (DGHO) ist das PFS ein sinnvoller und gebräuchlicher Endpunkt bei allen Therapieinterventionen mit palliativer Zielsetzung. Gegenüber dem Endpunkt Überlebenszeit hat der Endpunkt PFS den Vorteil einer geringeren Beeinflussung durch spätere Störfaktoren wie Cross-over und / oder Folgetherapien.

Ergänzend zu PFS wird in dem vorliegenden Dossier die ORR berichtet. Die ORR wurde definiert als der Anteil der Patienten mit dem besten Gesamtansprechen (engl. Best Overall Response, BOR) in Form eines vollständigen Tumoransprechens (engl. Complete Response, CR) oder eines teilweisen Tumoransprechens (engl. Partial Response, PR). Die ORR stellt somit ein Maß für die Kontrolle des Tumorwachstums dar (Eisai Inc. 2017b).

Ein Ansprechen auf die Therapie wurde, analog zum PFS, unter Verwendung von IIR gemäß mRECIST bzw. als Sensitivitätsanalyse gemäß RECIST v.1.1 sowie durch den Prüfarzt gemäß mRECIST gemessen. Unter CR versteht man dabei gemäß mRECIST den vollständigen Rückgang aller intratumoralen arteriellen Erweiterungen der Zielläsionen (Lencioni und Llovet 2010) und unter PR eine Reduktion der Summe der Durchmesser der lebensfähigen Zielläsionen um mindestens 30% im Vergleich zum Ausgangswert (Lencioni und Llovet 2010). Gemäß RECIST v.1.1 ist CR definiert als der vollständige Rückgang aller Zielläsionen. PR ist hingegen definiert als die Reduktion der Summe der Durchmesser der Zielläsionen um mindestens 30% im Vergleich zum Ausgangswert (Lencioni und Llovet 2010).

Die ORR wird in diesem Dossier als ergänzende Analyse berichtet, aber nicht zur Ableitung des Zusatznutzens herangezogen.

EORTC QLQ-C30

Der Fragebogen EORTC QLQ-C30 ist in der Onkologie ein häufig eingesetztes Instrument zur Bewertung der Morbidität und Lebensqualität mit insgesamt 30 Fragen. Die gestellten Fragen beziehen sich dabei auf die Symptome Fatigue, Übelkeit und Erbrechen, Schmerz, Dyspnoe, Schlaflosigkeit, Appetitminderung, Obstipation, Diarrhö sowie auf die Funktionen Physische Funktion, Rollenfunktionsniveau, Emotionale Funktion, Kognitive Funktion, Soziale Funktion, Finanzielle Schwierigkeiten und allgemein auf den Globalen Gesundheitsstatus. Die einzelnen Items werden dabei auf einer Skala von 1 bis 4 bzw. 1 bis 7, für die Fragen zum globalen Gesundheitsstatus, bewertet, die zur Auswertung in eine Skala von 0 bis 100 transformiert wird. Höhere Werte auf den Funktionsskalen und der Globalskala bedeuten eine höhere Lebensqualität und auf den Symptomskalen eine Verschlechterung der Symptome (Giesinger et al. 2016). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen (engl. minimal important difference, MID) (Cocks et al. 2011; Osoba et al. 1998).

Der EORTC QLQ-C30 ist eines der höchstentwickelten und am häufigsten verwendeten Messinstrumente zur Erhebung der Lebensqualität bei Patienten mit einer Krebserkrankung. Die Validität des Patientenfragebogens wurde für verschiedene onkologische Indikationen bestätigt (Aaronson et al. 1993). Darüber hinaus ist der EORTC QLQ-C30 als validiertes

Messinstrument gemäß Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) anerkannt (G-BA 2014; IQWiG 2013, 2014; Medac 2013).

Im Rahmen der vorliegenden Bewertung werden die Funktionsskalen zur Erfassung der gesundheitsbezogenen Lebensqualität herangezogen, während die Symptomskalen der Messung der Krankheitssymptome dienen.

EORTC QLQ-HCC18

Der EORTC QLQ-HCC18 ist ein HCC-spezifischer Fragebogen mit 18 Fragen bezüglich der Symptome Fatigue, Ikterus, Schmerz, Fieber und Abdominale Schwellung sowie der Items Körperbild, Ernährung und Sexualleben und wurde entwickelt, um die Sensitivität und Spezifität von HCC-bezogenen Fragestellungen im Hinblick auf die gesundheitsbezogene Lebensqualität zu erhöhen. Die einzelnen Items werden auf einer Skala von 1 bis 4 bewertet, welche wiederum linear in eine Skala von 0 bis 100 umgewandelt wird. Der Wert 100 stellt dabei den schlechtesten Status dar (Chie et al. 2012). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen.

Für die Bewertung der Morbidität wurden die Symptomskalen zu Fatigue, Ikterus, Schmerz, Fieber, Abdominale Schwellung und für die Bewertung der gesundheitsbezogenen Lebensqualität die Items Körperbild, Ernährung und Sexualleben des EORTC QLQ-HCC18 herangezogen.

Sowohl die Ergebnisse des EORTC QLQ-C30 als auch die Ergebnisse des EORTC QLQ-HCC18 sind hochrelevante prognostische Faktoren für das Gesamtüberleben bei neu diagnostizierten HCC-Patienten (Li et al. 2017). Das hierbei verwendetet Scoring-Verfahren ist eine effektive Möglichkeit zur Zusammenfassung, Analyse und Interpretation von Rohdaten zur Morbidität und gesundheitsbezogenen Lebensqualität und macht die beiden Fragebögen EORTC QLQ-C30 und EORTC QLQ-HCC18 in der klinischen Praxis aussagekräftig und kommunikationsfähig (Chie et al. 2012; Li et al. 2017).

EQ-5D VAS

Der Fragebogen EQ-5D ist ein von der internationalen Forschungsgruppe EuroQoL entwickeltes, verlässliches und validiertes Messinstrument für den Gesundheitszustand eines Patienten. Er besteht aus einem deskriptiven Teil und einer visuellen Analogskala (engl. Visual Analogue Scale, VAS), die zusätzlich in den Fragebogen aufgenommen wurde, um den aktuellen Gesundheitszustand auf einer Skala von 0 bis 100 zu messen, wobei 0 für den "schlechtesten denkbaren Gesundheitszustand" und 100 für den "besten denkbaren Gesundheitszustand" steht. (Pickard et al. 2012). Der Fragebogen wurde und wird derzeit in zahlreichen Studien mit Krebspatienten sowie speziell bei Patienten mit HCC eingesetzt (Pickard et al. 2012; Teckle et al. 2013). Eine Veränderung auf der VAS eines Patienten um mindestens 7 oder 10 mm im Vergleich zum Ausgangswert wurde hierbei als Response definiert (Pickard et al. 2007).

Die EQ-5D VAS stellt somit ein allgemeines quantitatives Maß für den Gesundheitszustand eines Patienten entsprechend seiner eigenen Beurteilung dar. Im Sinne der Berücksichtigung der Lebensqualität und Symptomatik der Patienten im Rahmen der Nutzenbewertung ist die EQ-5D VAS als patientenrelevanter Endpunkt anzusehen (IQWiG 2017).

Gesundheitsbezogene Lebensqualität

Die gesundheitsbezogene Lebensqualität ist eine wichtige Zielgröße bei der medizinischen Versorgung krebskranker Patienten (IQWiG 2011). Sie spiegelt direkt die Auswirkungen der Erkrankung auf das Befinden des Patienten und seinen Alltag wider und ist damit unmittelbar patientenrelevant. Auch der Erhalt der Lebensqualität stellt gerade in der palliativen Situation, die auf Dauer fast immer mit einer Verschlechterung der Lebensqualität einhergeht, einen patientenrelevanten Endpunkt dar (IQWiG 2017).

Die gesundheitsbezogene Lebensqualität wurde in der eingeschlossenen Studie 304 anhand der Funktionsskalen Physische Funktion, Rollenfunktionsniveau, Emotionale Funktion, Kognitive Funktion, Soziale Funktion, Finanzielle Schwierigkeiten und Globaler Gesundheitsstatus des Fragebogens EORTC QLQ-C30 beurteilt. Außerdem wurden für die Bewertung der gesundheitsbezogenen Lebensqualität die Items Körperbild, Ernährung und Sexualleben des HCC-spezifischen EORTC QLQ-HCC18 herangezogen.

Sicherheit und Verträglichkeit

Im vorliegenden Dossier werden die Kategorien UE, SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 ausführlich dargestellt. UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, werden ergänzend dargestellt.

Die Kodierung der UE erfolgte unter Verwendung des Medical Dictionary for Regulatory Activities (MedDRA) Version 19.1. Die Einteilung in verschiedene Grade erfolgte nach CTCAE Version 4.0.

Alle UE und SUE wurden als Ereignisse erfasst, die während der Behandlung auftraten (engl. Treatment Emergent Adverse Events, TEAE) erfasst. TEAE wurden definiert als jedes Ereignis,

- das zu Baseline nicht präsent war und während oder bis zu 30 Tage nach der Behandlung mit der Studienmedikation auftrat.
- das bei einer Vorbehandlung aufgetreten ist, aber zu Baseline nicht mehr präsent war und während oder bis zu 30 Tage nach der Behandlung mit der Studienmedikation wieder auftrat.
- dessen Schweregrad sich während oder bis zu 30 Tage nach der Behandlung im Vergleich zum Zustand vor der Behandlung mit der Studienmedikation verschlechterte.

Die im Rahmen dieser Nutzenbewertung berichteten UE zählen zu den Aspekten der Sicherheit und Verträglichkeit und sind somit vom IQWiG anerkannte patientenrelevante Endpunkte (IQWiG 2017).

Statistische Methoden zur Bewertung der primären und sekundären Endpunkte

Die Wirksamkeitsanalysen von OS und PFS wurden auf Grundlage der ITT-Population gemäß dem ITT-Prinzip durchgeführt. Die ITT-Population umfasst dabei alle Patienten, die randomisiert wurden. Die Gruppenzuteilung für die Analyse erfolgt anhand der geplanten Gruppenzuteilung unabhängig von der tatsächlichen Behandlung. Die Analysen der Fragebögen EORTC QLQ-C30, EORTC QLQ-HCC18 und EQ-5D wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. Die Analysen zu Sicherheit und Verträglichkeit basierten auf dem Safety Analysis Set (SAS). Das SAS umfasst alle Patienten, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Für die Sicherheitsanalysen erfolgte die Gruppenzuteilung für die Analyse anhand der tatsächlichen Behandlung unabhängig von der geplanten Gruppenzuteilung.

Für den Vergleich von Lenvatinib und Sorafenib wurden Analysen mit stratifizierten Modellen, stratifiziert nach den Stratifizierungsfaktoren gemäß interaktivem Voice and Web Response System (IxRS), durchgeführt. Dabei handelte es sich um folgende Stratifizierungsfaktoren:

- Region („Asien-Pazifik“ vs. „westliche Regionen“)
- MPVI oder ES oder beides („Ja“ vs. „Nein“)
- Baseline Eastern Cooperative Oncology Group-Performance Status (ECOG-PS) („0“ vs. „1“)
- Baseline Gewichtsgruppe („< 60 kg“ vs. „≥ 60 kg“)

Während des Studienverlaufs gab es zwei im Voraus geplante und im statistischen Analyseplan (SAP) definierte Zwischenanalysen. Die erste Zwischenanalyse wurde am 13. Mai 2015 durchgeführt, nachdem etwa 30% der notwendigen Fallzahl (entsprach 249 Todesfällen) in beiden Behandlungsarmen beobachtet wurde. Die zweite Zwischenanalyse wurde am 03. Februar 2016 durchgeführt, nachdem circa 70% der notwendigen Fallzahl (entsprach 521 Todesfällen) in beiden Behandlungsarmen beobachtet wurde. Basierend auf beiden von einem unabhängigen Statistikerteam durchgeführten Zwischenanalysen wurde vom Data Safety Monitoring Board (DSMB) empfohlen, die Studie ohne weitere Modifikationen fortzuführen. Die finale Analyse erfolgte am 13.November 2016.

Gesamtüberleben (OS)

Für den Endpunkt OS wurde als primäre Analyse entsprechend dem SAP ein Nicht-Unterlegenheits-Test von Lenvatinib gegenüber Sorafenib durchgeführt. Dazu wurden das Hazard Ratio (HR) und das dazugehörige 95%-Konfidenzintervall (KI) mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS

Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Die Efron-Methode wurde zur Korrektur von Bindungen verwendet.

Für das OS wurde zudem ein Überlegenheits-Test von Lenvatinib gegenüber Sorafenib durchgeführt. Dazu wurde ein zweiseitiger stratifizierter Log-Rank-Test, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 angewendet.

Da die Studie nicht aufgrund einer der Zwischenanalysen gestoppt wurde, wurde die Nicht-Unterlegenheit bezüglich OS erstmals bei der finalen Analyse getestet. Nicht-Unterlegenheit galt als erfüllt, wenn bei der finalen Analyse die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag. Falls Nicht-Unterlegenheit für den Endpunkt OS gezeigt wurde, sollte auf Überlegenheit getestet werden. Überlegenheit galt als erfüllt, wenn mithilfe des stratifizierten Log-Rank-Tests bei der finalen Analyse der zweiseitige p-Wert für den Test auf Gleichheit < 0,05 betrug.

Das mediane OS wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das zum medianen OS gehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Überlebenswahrscheinlichkeiten wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt.

Progressionsfreies Überleben (PFS)

Für den Endpunkt PFS wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet.

Der Unterschied im PFS zwischen Lenvatinib und Sorafenib wurde mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert.

Das mediane PFS wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das zum medianen PFS gehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für ein progressionsfreies Überleben wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt.

Für die Ableitung des Zusatznutzens werden in diesem Dossier die tumorbezogenen Analysen des PFS basierend auf den Daten des IIR per mRECIST als Hauptanalysen herangezogen.

Ergänzend werden Analysen zur ORR dargestellt, um die Ergebnisse zum PFS zu untermauern. Dazu wurde die ORR (basierend auf den mRECIST IIR Daten) für beide Behandlungsarme bestimmt und der Unterschied zwischen Lenvatinib und Sorafenib mittels Odds Ratio (OR), dem relativen Risiko (engl. Risk Ratio, RR) und der Risikodifferenz (engl. Risk Difference, RD) und den dazugehörigen 95%-KIs analysiert. Die OR und die statistische Signifikanz des

Unterschieds in der ORR zwischen Lenvatinib und Sorafenib wurde mittels stratifiziertem Cochran-Mantel-Haenszel Chi-Quadrat-Test, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 untersucht. Bei der Analyse der ORR handelt es sich um eine ergänzende Analyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

EORTC QLQ-C30

Um die Validität der Auswertungen des EORTC QLQ-C30-Fragebogens bezüglich der Symptome Fatigue, Übelkeit und Erbrechen, Schmerz, Dyspnoe, Schlaflosigkeit, Appetitminderung, Obstipation und Diarrhö sowie der Funktionen Physische Funktion, Rollenfunktionsniveau, Emotionale Funktion, Kognitive Funktion, Soziale Funktion, Finanzielle Schwierigkeiten und Globaler Gesundheitsstatus zu zeigen, wurden die entsprechenden Rücklaufquoten angegeben.

Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms bzw. der jeweiligen Funktion wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet.

Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms bzw. der jeweiligen Funktion wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms bzw. der jeweiligen Funktion wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt.

EORTC QLQ-HCC18

Um die Validität der Auswertungen des EORTC QLQ-HCC18-Fragebogens bezüglich der Items Fatigue, Ikterus, Schmerz, Fieber, Abdominale Schwellung, Körperbild, Ernährung und Sexualleben zu zeigen, wurden die entsprechenden Rücklaufquoten angegeben.

Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet.

Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert.

Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt.

EQ-5D VAS

Um die Validität der Auswertungen des EQ-5D-Fragebogens bezüglich der Visual Analogue Scale (VAS) zu zeigen, wurde die entsprechende Rücklaufquote angegeben.

Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet.

Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert.

Die mediane Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt.

Sicherheit und Verträglichkeit

Da die Gesamtbeobachtungszeiten für Lenvatinib deutlich länger waren als in der Kontrollgruppe Sorafenib, wären die Inzidenzraten und die darauf basierenden Effektmaße RR, OR und RD zuungunsten von Lenvatinib verzerrt, so dass nur sehr eingeschränkt Aussagen zu einem größeren oder vergleichbaren bzw. geringeren Schaden gemacht werden könnten. Daher wurden für die Hauptkategorien zu Sicherheit und Verträglichkeit UE, SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 sowie für die ergänzenden Endpunkte UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, Time-To-First-Event-Analysen durchgeführt. Das HR und das dazugehörige 95%-KI wurden mithilfe des unstratifizierten Cox-Proportional-Hazard-Modells berechnet. Auf die Stratifizierung wurde verzichtet, um einerseits für alle Sicherheitsendpunkte dieselbe Analysemethode zu verwenden und andererseits zu berücksichtigen, dass insbesondere

für die UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, teils sehr wenige Ereignisse berichtet wurden und dadurch die Konvergenz der statistischen Modelle nicht gewährleistet wäre, wenn die Strata berücksichtigt worden wären.

Die mediane Zeit bis zum ersten Ereignis wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das zur medianen Zeit gehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten[4] erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet[5] werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität[6] erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten

[4] *DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-188.*
[5] *Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.*
[6] *Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.*

Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Eine Meta-Analyse wurde nicht durchgeführt.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Im Rahmen des Dossiers werden folgende Sensitivitätsanalysen auf Endpunktebene durchgeführt und berichtet, um die Aussagen der Hauptanalysen und ihre Ergebnissicherheit in einem umfassenden Kontext darzustellen.

Gesamtüberleben (OS)

Aufgrund von Imbalancen in den Patientencharakteristika zu Baseline zwischen den beiden Studienarmen (siehe Abschnitt 4.3.1.2.1) wurden uni- und multivariate adjustierte Analysen post-hoc durchgeführt. Für die Bestimmung von für das HCC prognostischen Patientencharakteristika Charakteristika, wurden Mediziner der Europäischen Union befragt, die Patienten mit inoperablem HCC behandeln. Die Auswahl der Mediziner mit relevantem Expertenwissen und Erfahrung, die in die Befragung einwilligten, erfolgte anhand eines nicht-probabilistischen Verfahrens. Insgesamt wurden fünf Mediziner via E-Mail kontaktiert, um ihre Expertenmeinung bezüglich der Relevanz von Patientencharakteristika als prognostische Faktoren für HCC einzuholen. Die Basis bilden dabei die Patientencharakteristika, die in der Studie 304 erfasst wurden. Für alle von den Medizinern als relevant identifizierten Patientencharakteristika wurden univariate Analysen mittels Cox-Proportional-Hazard-Modell durchgeführt und der Einfluss auf das HR analysiert. Zusätzlich wurde eine multivariate Analyse mittels Cox-Proportional-Hazard-Modell durchgeführt, wobei anhand einer vorwärts gerichteten schrittweisen Methode die Kovariaten gewählt wurden, die die Daten bezüglich des Wald-Tests, des Likelihood-Verhältnis-Tests und des Aikake’s Informationskriteriums am

besten abbilden. Die uni- und multivariaten Analysen werden zur Ableitung des Zusatznutzens herangezogen (Eisai Inc. 2018a).

Progressionsfreies Überleben (PFS)

Die Robustheit der Ergebnisse der Hauptanalyse des Endpunkts PFS, bestimmt mittels IIR per mRECIST, wird durch die tumorbezogenen Analysen des PFS basierend auf den Auswertungen durch den Prüfarzt per mRECIST und basierend auf den Daten des IIR per RECIST v.1.1. untermauert. Hierbei handelt es sich um Sensitivitätsanalysen, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

Ergänzend werden Analysen der ORR dargestellt, um die Ergebnisse des PFS zu stützen. Die Robustheit des Ergebnisses der Analyse des Endpunkts ORR, bestimmt mittels IIR per mRECIST, wird durch die tumorbezogenen Analysen des ORR basierend auf den Auswertungen durch den Prüfarzt per mRECIST und basierend auf den Daten des IIR per RECIST v.1.1 untermauert. Diese Sensitivitätsanalysen der ergänzenden Analyse der ORR werden nicht für die Ableitung des Zusatznutzens herangezogen.

EORTC QLQ-C30

Die Robustheit der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items, gemessen auf den Skalen des EORTC QLQ-C30, wird mit Hilfe der mittleren Veränderung zu Baseline im Score untermauert. Dazu wurden der Mittelwert zu Baseline und der Least Square (LS)-Mittelwert der Änderung zu Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zu Baseline analysiert. Bei der Analyse der mittleren Veränderung zu Baseline im Score des jeweiligen Items handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

EORTC QLQ-HCC18

Die Robustheit der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items, gemessen auf den Skalen des EORTC QLQ-HCC18, wird mit Hilfe der mittleren Veränderung zu Baseline im Score untermauert. Dazu wurden der Mittelwert zu Baseline und der LS-Mittelwert der Änderung zu Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zu Baseline analysiert. Bei der Analyse der mittleren Veränderung zu Baseline im Score des jeweiligen Items handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

EQ-5D VAS

Die Robustheit der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wird mit Hilfe der mittleren Veränderung zu Baseline auf der VAS untermauert. Dazu wurden der Mittelwert zu Baseline und der LS-Mittelwert der Änderung zu Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zu Baseline analysiert.

Bei der Analyse der mittleren Veränderung zu Baseline auf der VAS handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

– Geschlecht
– Alter
– Krankheitsschwere bzw. –stadium
– Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Folgende Subgruppenanalysen wurden für die Endpunkte OS und PFS, für die EORTC QLQ-C30 Symptom- und Funktionsskalen, für die Skalen des EORTC QLQ-HCC18, für die EQ-5D VAS sowie für die Hauptkategorien der Endpunkte zu Sicherheit und Verträglichkeit durchgeführt.

Tabelle 4-F: Subgruppenanalysen der Studie 304

Kategorie	**Ausprägung**
Baseline Altersgruppe (kategorial)	▪ < 65 Jahre ▪ ≥ 65 bis <75 Jahre ▪ ≥ 75 Jahre
Geschlecht (kategorial)	▪ Männlich ▪ Weiblich
Region[a] (kategorial)	▪ Asien-Pazifik ▪ Westliche Regionen
Makroskopische Invasion der Pfortader (MPVI) (kategorial)	▪ Ja ▪ Nein
Extrahepatische Ausbreitung (ES) (kategorial)	▪ Ja ▪ Nein
MPVI oder ES oder beides (kategorial)	▪ Ja ▪ Nein
Baseline ECOG-PS (kategorial)	▪ 0 ▪ 1
Baseline Gewichtsgruppe (kategorial)	▪ < 60 kg ▪ ≥ 60 kg
Baseline AFP (kategorial)	▪ < 200 ng/mL ▪ ≥ 200 ng/mL
Ätiologie HBV (kategorial)	▪ Ja ▪ Nein
Ätiologie HCV (kategorial)	▪ Ja ▪ Nein
Ätiologie Alkohol (kategorial)	▪ Ja ▪ Nein
BCLC Staging (kategorial)	▪ Stadium B ▪ Stadium C
Leberzirrhose zu Baseline (CRF) (kategorial)	▪ Ja ▪ Nein
Leberzirrhose zu Baseline (IIR) (kategorial)	▪ Ja ▪ Nein
Vorbehandlung mit TACE (kategorial)	▪ Ja ▪ Nein

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die Asien-Pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); kg: Kilogramm; HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); ng: Nanogramm; mL: Milliliter; TACE: Transarterielle Chemoembolisation

Die post-hoc definierten Subgruppenkategorien Leberzirrhose zu Baseline auf Basis der Case Report Form (CRF; entspricht der zu Studienbeginn dokumentierten Anzahl an Patienten mit Leberzirrhose), Leberzirrhose zu Baseline auf Basis des IIR (entspricht der auf Basis der unabhängigen Bewertung radiologischer Verfahren dokumentierten Anzahl an Patienten mit Leberzirrhose), und Vorbehandlung mit transarterieller (Chemo)-Embolisation (TACE oder TAE) bei Patienten mit BCLC Stadium B wurden gemäß den im Rahmen der Beratung gegebenen Hinweisen des G-BA durchgeführt (G-BA 2018a) und auf Effektmodifikation hin untersucht.

Durchführung der Subgruppenanalysen

Sowohl für die in Abschnitt 4.2.5.4 beschriebenen Sensitivitätsanalysen als auch für die in Abschnitt 4.2.5.2 beschriebenen ergänzenden Analysen wurden keine Subgruppenanalysen durchgeführt.

Für alle Ereigniszeitanalysen wurden für jeden Endpunkt die Interaktions-p-Werte der Subgruppen auf der Basis eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells für proportionale Hazards, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, berechnet. Auf eine stratifizierte Analyse wurde verzichtet, da es in den Subgruppen zum Teil sehr wenige Ereignisse gab und mit Berücksichtigung der Strata in den Analysen die Konvergenz der statistischen Modelle nicht sichergestellt ist.

Die Subgruppenanalysen der Endpunkte OS (Mortalität) und PFS (Morbidität) wurden auf Basis der Daten der ITT-Population durchgeführt. Die Subgruppenanalyen der Endpunkte zu Morbidität und gesundheitsbezogenen Lebensqualität, die auf der Grundlage der Fragebögen EORTC QLQ-C30, EORTC QLQ-HCC18 und EQ-5D erhoben wurden, wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. Die Subgruppenanalysen der Endpunkte zu Sicherheit und Verträglichkeit basierten auf dem SAS.

Darstellung der Subgruppenanalysen

Die Ergebnisse des Interaktionstests werden im vorliegenden Dossier für jeden Endpunkt separat dargestellt. Dazu werden die Interaktions-p-Werte der Subgruppenmerkmale tabellarisch aufgeführt. Gemäß den allgemeinen Methoden des IQWiG wurde ein Interaktions-p-Wert von $< 0,05$ als Beleg für eine Interaktion zwischen Behandlung und Subgruppenmerkmal interpretiert (IQWiG 2017). Bei Vorliegen eines Belegs für eine Interaktion werden die Ergebnisse der einzelnen Subgruppen im nächsten Schritt ausführlich präsentiert. Die vollständigen Subgruppenanalysen sind im Modul 5 angegeben.

Interpretation der Subgruppenanalysen

Bei der Interpretation der Ergebnisse muss berücksichtigt werden, dass die Durchführung von Interaktionstests mit einer Irrtumswahrscheinlichkeit von 5% (bei einem p-Wert von 0,05) behaftet ist. Bei der Analyse zahlreicher Subgruppen besteht damit eine hohe Wahrscheinlichkeit, dass die Ergebnisse irgendeiner Subgruppe statistische Signifikanz

erreichen, obwohl es sich in Wahrheit um ein zufälliges Ergebnis handelt (IQWiG 2017). Aus diesem Grund und dem Umstand, dass die betrachteten Studien nicht für Subgruppenanalysen gepowert waren, ergibt sich die Notwendigkeit, Aussagen über die Effekte in den Subgruppen vorsichtig und keinesfalls konfirmatorisch zu interpretieren.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich[7]. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“[8], „Multiple-Treatment-Meta-Analysen“[9] oder auch „Netzwerk-Meta-Analysen“[10], sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist[11].

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vorliegen.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen[12].

---

[7] *Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-148.*

[8] *Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.*

[9] *Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005; 331(7521): 897-900.*

[10] *Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. Stat Methods Med Res 2008;17(3): 279-301.*

[11] *B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

[12] *Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009; 338: b1147.*

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*

- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*

- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*

- *Art und Umfang von Sensitivitätsanalysen.*

Es wurde kein indirekter Vergleich durchgeführt.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

– Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.3.1)
– Nicht randomisierte vergleichende Studien (Abschnitt 4.3.3.2)
– Weitere Untersuchungen (Abschnitt 4.3.3.3)

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*

- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungsstudie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer**	**Therapiearme**
E7080-G000-304 (NCT01761266)	ja	ja	abgeschlossen	193 Wochen	Lenvatinib Sorafenib

Geben Sie an, welchen Stand die Information in Tabelle 4-1 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information: 01. August 2018 (Eisai 2018b).

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-1 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
Nicht zutreffend	

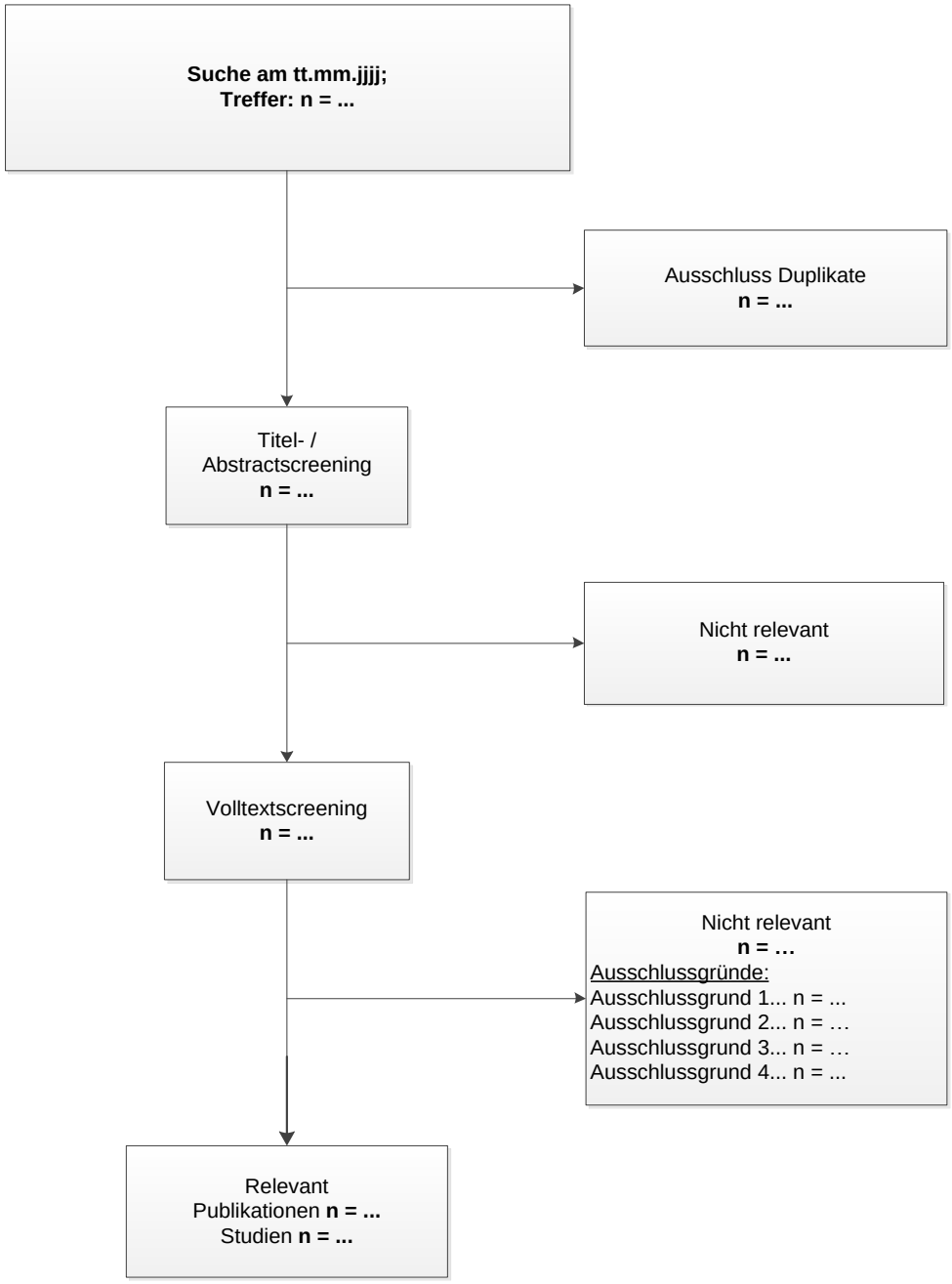
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.



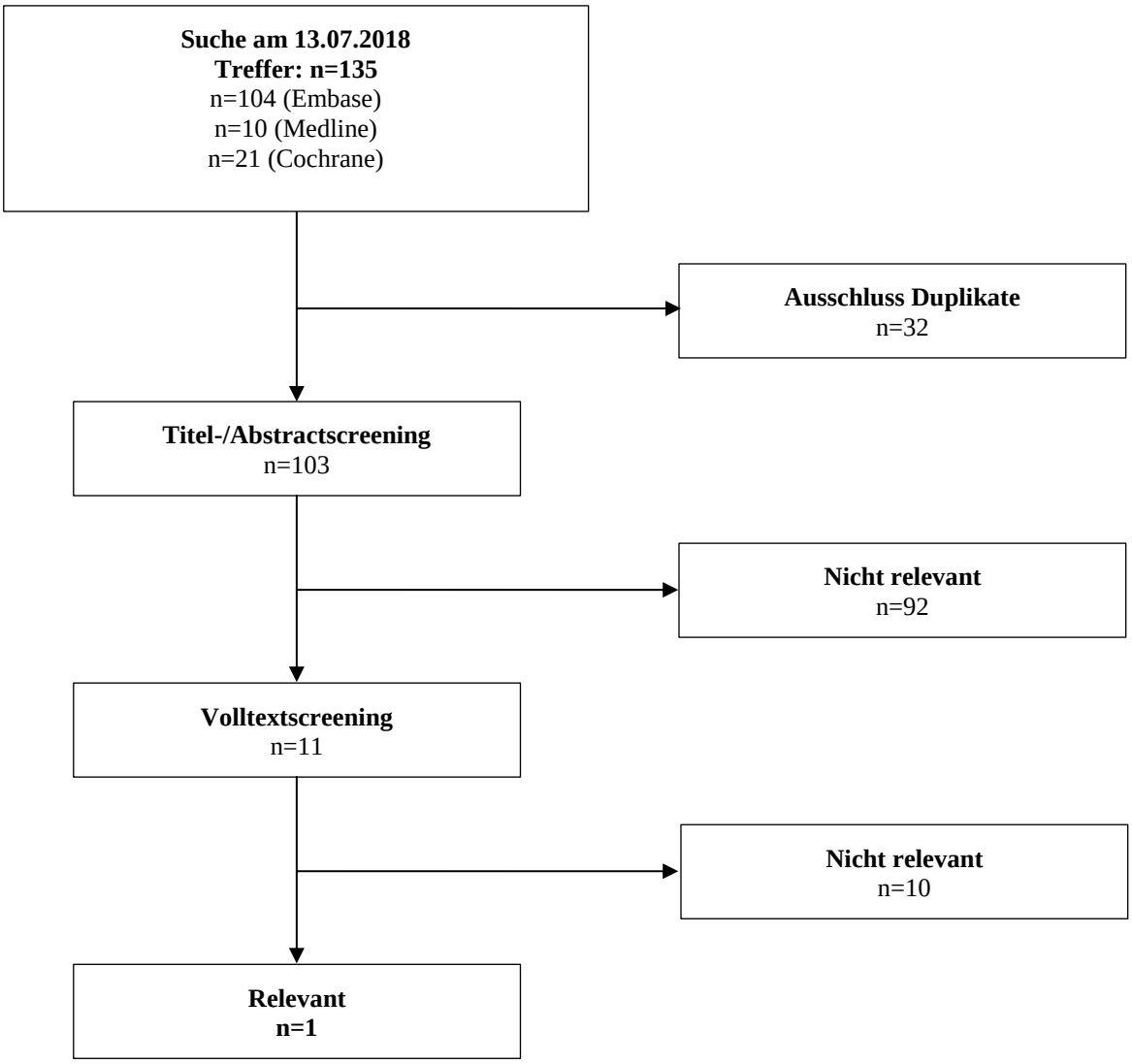


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-1) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	**Identifikationsorte (Name des Studienregisters und Angabe der Zitate[a])**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Status (abgeschlossen/ abgebrochen/ laufend)**
E7080-G000-304	Clinicaltrials.gov NCT00946153 (ClinicalTrialsGOV 2018) EU CTR 2012-002992-33 (EU-Clinical Trials Register 2012) PharmNet.Bund 2012-002992-33 (PharmnetBund 2012) WHO ICTRP NCT01761266 (WHO ICTRP 2018)	ja	ja	abgeschlossen
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. EU CTR: European Union Clinical Trails Registery; ICTRP: Internation Clinical Trials Registry Platform Search Portal; WHO: World Health Organization				

Geben Sie an, welchen Stand die Information in Tabelle 4-3 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information: 06. Juli 2018.

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen[a]		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie[b] (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag[c] (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
E7080-G000-304	ja	ja	nein	ja (Eisai Inc. 2017b)	Ja Clinicaltrials.gov NCT00946153 (ClinicalTrialsGOV 2018) EU CTR 2012-002992-33 (EU-Clinical Trials Register 2012) PharmNet.Bund 2012-002992-33 (PharmnetBund 2012) WHO ICTRP NCT01761266 (WHO ICTRP 2018)	Ja (Kudo et al. 2018)

a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.

b: Studie, für die der Unternehmer Sponsor war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – Studie 304 mit dem zu bewertenden Arzneimittel

Studie	**Studiendesign** <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)[a]**	**Studiendauer** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
304	RCT, multizentrisch, 1:1 Randomisierung, Open-Label, Phase-III Nicht-Unterlegenheits-Studie	Erwachsene Patienten mit histologisch oder zytologisch bestätigtem inoperablem HCC oder nach AASLD Kriterien klinisch bestätigtem HCC, einschließlich Leberzirrhose jeglicher Ätiologie, oder mit chronischer HBV oder HCV Infektion	Lenvatinib (N=478) Sorafenib (N=476)	1. Prä-Randomisierungsphase (bestehend aus Screening- und Baseline-Phase): bis zu 21 Tage 2. Randomisierungsphase (bestehend aus Behandlungs- und Nachbeobachtungsphase): Ab Randomisierung des ersten Patienten bis zum Zeitpunkt des Datenschnitts (Eintritt von insgesamt 700 Todesfällen über beide Behandlungsarme) 3. Extensionsphase (bestehend aus Behandlungs- und Nachbeobachtungsphase[a])	183 Zentren in Asien, Nordamerika, der EU, Russland, Israel Zeitraum der Durchführung: 01.03.2013 bis 13.11.2016 (Datenschnitt für die Analyse des primären Endpunkts)	Primärer Endpunkt: OS Patientenrelevante sekundäre Endpunkte: PFS EORTC QLQ-C30 EORTC QLQ-HCC18 EQ-5D VAS Sicherheit und Verträglichkeit

a: Alle Patienten, die nach dem Datenschnitt für die Analyse des primären Endpunkts weiterhin die Studienmedikation erhielten oder in der Nachbeobachtungsphase waren, sind in die Extensionsphase überführt und dort wie in der Randomisierungsphase weiterbehandelt oder nach Therapiebeendigung alle 12 Wochen bis zum Tod weiterbeobachtet worden, um das OS und nachfolgende Krebsbehandlungen zu dokumentieren.

AASLD: American Association for the Study of Liver Diseases; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; EU: Europäisch Union; HBV: Hepatitis B Virus; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HCV: Hepatitis C Virus; N: Anzahl der Patienten; OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); RCT: Randomized Controlled Trial (randomisierte kontrollierte Studie)

Quelle: Eisai Inc. 2017b

Tabelle 4-6: Charakterisierung der Interventionen – Studie 304 mit dem zu bewertenden Arzneimittel

Studie	**Lenvatinib**	**Sorafenib**	***ggf. weitere Spalten mit Behandlungscharakteristika*** ***z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.***
304	Patienten mit Körpergewicht < 60 kg: 8 mg (zwei 4 mg-Kapseln), Einnahme einmal täglich, oral Patienten mit einem Körpergewicht ≥ 60 kg: 12 mg (drei 4 mg-Kapseln), Einnahme einmal täglich, oral	Tagesdosis: 800 mg, Einnahme zweimal täglich zwei 200 mg Tabletten, oral	Die Behandlung wurde aufgrund einer Krankheitsprogression oder inakzeptabler Toxizität, auf Wunsch des Patienten oder aufgrund des Widerrufs der Einwilligung des Patienten abgebrochen
kg: Kilogramm; mg: Milligramm Quelle: Eisai Inc. 2017b			

Nach der Beschreibung der Interventionen der Studie 304 ist in der nachfolgenden Tabelle die Studienpopulation charakterisiert (Tabelle 4-7:). Da Lenvatinib in Abhängigkeit vom Körpergewicht zu Baseline (< 60 kg und ≥ 60 kg) in unterschiedlichen Dosen (8 mg und 12 mg) verabreicht wurde, erfolgt auch die Charakterisierung der Studienpopulation separat für 8 mg und 12 mg der Studienmedikation Lenvatinib.

Tabelle 4-7: Charakterisierung der Studienpopulationen: demographische Charakteristika zu Studienbeginn – Studie 304 mit dem zu bewertenden Arzneimittel

Charakteristika	**Lenvatinib**			**Sorafenib (N=476)**	**Gesamt (N=954)**
	8 mg[a] (N=151)	**12 mg[a] (N=327)**	**Gesamt (N=478)**		
Baseline Alter (Jahre)					
Mittelwert (SD)	63,1 (12,30)	60,4 (11,32)	61,3 (11,69)	61,2 (12,01)	61,3 (11,84)
Median (Min; Max)	65,0 (20; 86)	62,0 (24; 88)	63,0 (20; 88)	62,0 (22; 88)	62,0 (20; 88)
Baseline Altersgruppe (Jahre) (N (%))					
< 65	69 (45,7)	201 (61,5)	270 (56,5)	283 (59,5)	553 (58,0)
≥ 65 bis < 75	56 (37,1)	94 (28,7)	150 (31,4)	126 (26,5)	276 (28,9)
≥ 75	26 (17,2)	32 (9,8)	58 (12,1)	67 (14,1)	125 (13,1)
Geschlecht (N (%))					
Männlich	106 (70,2)	299 (91,4)	405 (84,7)	401 (84,2)	806 (84,5)
Weiblich	45 (29,8)	28 (8,6)	73 (15,3)	75 (15,8)	148 (15,5)
Region[b] (N (%))					
Asien-Pazifik	130 (86,1)	191 (58,4)	321 (67,2)	319 (67,0)	640 (67,1)
Westliche Regionen	21 (13,9)	136 (41,6)	157 (32,8)	157 (33,0)	314 (32,9)
Ethnische Gruppe (N (%))					
Weiße	17 (11,3)	118 (36,1)	135 (28,2)	141 (29,6)	276 (28,9)
Schwarze oder Afroamerikaner	0 (0,0)	7 (2,1)	7 (1,5)	6 (1,3)	13 (1,4)
Asiaten	134 (88,7)	200 (61,2)	334 (69,9)	326 (68,5)	660 (69,2)
Anteil Japaner	40 (26,5)	40 (12,2)	80 (16,7)	86 (18,1)	166 (17,4)
Anteil Chinesen	61 (40,4)	97 (29,7)	158 (33,1)	157 (33,0)	315 (33,0)
Anteil andere Asiaten	33 (21,9)	63 (19,3)	96 (20,1)	83 (17,4)	179 (18,8)
Indianer oder alaskanische Ureinwohner	0 (0,0)	1 (0,3)	1 (0,2)	0 (0,0)	1 (0,1)
Eingeborene Hawaiianer oder andere pazifische Inselbewohner	0 (0,0)	0 (0,0)	0 (0,0)	1 (0,2)	1 (0,1)
Andere	0 (0,0)	1 (0,3)	1 (0,2)	2 (0,4)	3 (0,3)
Ethnizität (N (%))					
Spanisch oder lateinamerikanisch	1 (0,7)	5 (1,5)	6 (1,3)	11 (2,3)	17 (1,8)
Nicht spanisch oder lateinamerikanisch	150 (99,3)	322 (98,5)	472 (98,7)	465 (97,7)	937 (98,2)

Charakteristika	Lenvatinib			Sorafenib (N=476)	Gesamt (N=954)
	8 mg[a] (N=151)	12 mg[a] (N=327)	Gesamt (N=478)		
Baseline Körpergewicht (kg)					
Mittelwert (SD)	52,7 (4,90)	75,9 (14,40)	68,6 (16,32)	68,1 (13,90)	68,3 (15,16)
Median (Min; Max)	53,0 (39; 60)	72,0 (60; 142)	66,2 (39; 142)	67,0 (39; 123)	66,9 (39; 142)
Baseline Gewichtsgruppe (kg) (N (%))					
<60 kg	151 (100,0)	2 (0,6)	153 (32,0)	146 (30,7)	299 (31,3)
≥60 kg	0 (0,0)	325 (99,4)	325 (68,0)	330 (69,3)	655 (68,7)
Baseline Körpergröße (cm)					
Mittelwert (SD)	160,7 (7,78)	170,1 (7,08)	167,1 (8,52)	167,2 (8,39)	167,2 (8,45)
Median (Min; Max)	161,0 (140; 178)	170,0 (147; 198)	167,6 (140; 198)	168,0 (140; 196)	168,0 (140; 198)
Baseline ECOG-PS (N (%))					
0	93 (61,6)	211 (64,5)	304 (63,6)	301 (63,2)	605 (63,4)
1	58 (38,4)	116 (35,5)	174 (36,4)	175 (36,8)	349 (36,6)
NYHA-Klassifikation (N (%))					
I	4 (2,6)	33 (10,1)	37 (7,7)	44 (9,2)	81 (8,5)
II	1 (0,7)	7 (2,1)	8 (1,7)	6 (1,3)	14 (1,5)
Nicht zutreffend	145 (96,0)	287 (87,8)	432 (90,4)	426 (89,5)	858 (89,9)
Unbekannt	1 (0,7)	0 (0,0)	1 (0,2)	0 (0,0)	1 (0,1)
Child-Pugh Score (N (%))					
5	111 (73,5)	257 (78,6)	368 (77,0)	357 (75,0)	725 (76,0)
6	40 (26,5)	67 (20,5)	107 (22,4)	114 (23,9)	221 (23,2)
7	0 (0,0)	3 (0,9)	3 (0,6)	4 (0,8)	7 (0,7)
8	0 (0,0)	0 (0,0)	0 (0,0)	1 (0,2)	1 (0,1)
MPVI (N (%))					
Ja	38 (25,2)	71 (21,7)	109 (22,8)	90 (18,9)	199 (20,9)
Nein	113 (74,8)	256 (78,3)	369 (77,2)	386 (81,1)	755 (79,1)
ES (N (%))					
Ja	91 (60,3)	200 (61,2)	291 (60,9)	295 (62,0)	586 (61,4)
Nein	60 (39,7)	127 (38,8)	187 (39,1)	181 (38,0)	368 (38,6)

Charakteristika	Lenvatinib			Sorafenib (N=476)	Gesamt (N=954)
	8 mg[a] (N=151)	12 mg[a] (N=327)	Gesamt (N=478)		
MPVI oder ES oder beides (N (%))					
Ja	105 (69,5)	224 (68,5)	329 (68,8)	336 (70,6)	665 (69,7)
Nein	46 (30,5)	103 (31,5)	149 (31,2)	140 (29,4)	289 (30,3)
Zugrundeliegende Leberzirrhose (N (%))[c]					
Ja	75 (49,7)	168 (51,4)	243 (50,8)	231 (48,5)	474 (49,7)
Nein	76 (50,3)	159 (48,6)	235 (49,2)	245 (51,5)	480 (50,3)

a: 8 mg und 12 mg waren die Start-Dosen basierend auf dem Gewicht des Probanden (< 60 kg und ≥ 60 kg)

b: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand

c: Der Anteil der Probanden mit zugrundeliegender Leberzirrhose (49,7%) wurde vermutlich unterschätzt, da diese Information im CRF über die Krankengeschichte aufgenommen und die Präsenz oder Abwesenheit einer Leberzirrhose nur verifiziert wurde, falls dies benötigt wurde, um die klinische Diagnose des HCC zu bestätigen.

cm: Zentimeter; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group Performance-Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); kg: Kilogramm; Max: Maximum; mg: Milligramm; Min: Minimum; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); NYHA: New York Heart Association; N: Anzahl der Patienten; SD: Standard Deviation (Standardabweichung)

Quelle: Eisai Inc. 2017b

Tabelle 4-8: Charakterisierung der Studienpopulationen: Krankheitsgeschichte zu Studienbeginn – Studie 304 mit dem zu bewertenden Arzneimittel

Charakteristika	**Lenvatinib (N=478)**	**Sorafenib (N=476)**	**Gesamt (N=954)**
Zeit seit erster Diagnose (Monate)			
Mittelwert (SD)	21,1 (30,17)	23,3 (34,66)	22,2 (32,49)
Median (Min; Max)	8,2 (0; 180)	9,0 (0; 250)	8,5 (0; 250)
Alter bei erster Diagnose (Jahre)			
Mittelwert (SD)	59,6 (11,57)	59,3 (11,54)	59,4 (11,55)
Median (Min; Max)	61,0 (15; 87)	60,0 (20; 85)	60,0 (15; 87)
BCLC Staging (N (%))			
Stadium B	104 (21,8)	92 (19,3)	196 (20,5)
Stadium C	374 (78,2)	384 (80,7)	758 (79,5)
Betroffene Krankheitsstellen (N (%))			
Leber	441 (92,3)	430 (90,3)	871 (91,3)
Lunge	163 (34,1)	144 (30,3)	307 (32,2)
Lymphknoten	127 (26,6)	141 (29,6)	268 (28,1)
Knochen	51 (10,7)	43 (9,0)	94 (9,9)
Andere	82 (17,2)	97 (20,4)	179 (18,8)
Anzahl der betroffenen Krankheitsstellen pro Patient (N (%))			
1	207 (43,3)	207 (43,5)	414 (43,4)
2	167 (34,9)	183 (38,4)	350 (36,7)
≥ 3	103 (21,5)	86 (18,1)	189 (19,8)
Faktor der Karzinogenese (N (%))			
HBV	251 (52,5)	228 (47,9)	479 (50,2)
HCV	91 (19,0)	126 (26,5)	217 (22,7)
Alkohol	36 (7,5)	21 (4,4)	57 (6,0)
Andere	38 (7,9)	32 (6,7)	70 (7,3)
Unbekannt	62 (13,0)	69 (14,5)	131 (13,7)
Leberzirrhose zu Baseline (N (%))[a]			
Ja	243 (50,8)	231 (48,5)	474 (49,7)
Nein	235 (49,2)	245 (51,5)	480 (50,3)
Baseline AFP (ng/mL)			

Charakteristika	**Lenvatinib (N=478)**	**Sorafenib (N=476)**	**Gesamt (N=954)**
Mittelwert (SD)	17507,5 (105137,39)	16678,5 (94789,46)	17096,5 (100088,76)
Median (Min; Max)	133,1 (0; 1567470)	71,2 (0; 1446396)	89,0 (0; 1567470)
Baseline AFP (N (%))			
< 200 ng/mL	255 (53,3)	286 (60,1)	541 (56,7)
≥ 200 ng/mL	222 (46,4)	187 (39,3)	409 (42,9)
Fehlt	1 (0,2)	3 (0,6)	4 (0,4)
Amoniakspiegel (µg/dL)			
Mittelwert (SD)	38,2 (29,98)	36,7 (32,90)	37,5 (31,46)
Median (Min; Max)	31,8 (4; 246)	30,0 (4; 473)	31,0 (4; 473)
Begleitende systemische antivirale Therapie von HBV oder HCV (N (%))			
Frühere Therapien zur Krebsbekämpfung	327 (68,4)	344 (72,3)	671 (70,3)
Strahlentherapie	49 (10,3)	60 (12,6)	109 (11,4)

a: Der Anteil der Probanden mit zugrundeliegender Leberzirrhose (49,7%) wurde vermutlich unterschätzt, da diese Information im CRF über die Krankengeschichte aufgenommen und die Präsenz oder Abwesenheit einer Leberzirrhose nur verifiziert wurde, falls dies benötigt wurde, um die klinische Diagnose des HCC zu bestätigen.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; dL: Deziliter; HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; Max: Maximum; Min: Minimum; µg: Mikrogramm; mL: Milliliter; N: Anzahl der Patienten; ng: Nanogramm; SD: Standard Deviation (Standardabweichung)

Quelle: Eisai Inc. 2017b

Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Für die Bewertung des Zusatznutzens von Lenvatinib wird die Studie 304 herangezogen. Diese wird im Folgenden hinsichtlich Studiendesign, Randomisierung, patientenrelevanter Endpunkte, Patientencharakteristika und Übertragbarkeit auf den deutschen Versorgungskontext beschrieben (Tabelle 4-5, Tabelle 4-6, Tabelle 4-7: , Tabelle 4-8).

Studiendesign

Bei der Studie 304 handelt es sich um eine multizentrische, randomisierte, Open-Label durchgeführte Nicht-Unterlegenheits- Phase-III-Studie, die die Wirksamkeit und Sicherheit von Lenvatinib zu Sorafenib in der systemischen Erstbehandlung von erwachsenen Patienten mit einem zytologisch bestätigten inoperablen HCC oder nach American Association for the

Study of Liver Diseases (AASLD) Kriterien klinisch bestätigtem HCC, einschließlich Leberzirrhose jeglicher Ätiologie, oder mit chronischer Hepatitis B Virus (HBV) oder Hepatitis C Virus (HCV) Infektion vergleicht.

Es waren 183 Zentren in Asien, Nordamerika, der EU, Russland und Israel beteiligt.

Die Studie selbst unterteilt sich in drei Studienabschnitte: 1. Prä-Randomisierungsphase, 2. Randomisierungsphase und 3. Extensionsphase. In der ersten Phase, der Prä-Randomisierungsphase, wurde zunächst die Eignung der Patienten für den Einschluss in die Studie festgestellt sowie die informierte Einwilligung zur Studienteilnahme eingeholt (Screeningphase). Das Screening erfolgte innerhalb von 21 Tagen vor der Randomisierung. Die Messungen der Baseline-Werte wurden zu Tag -1 vor Randomisierung oder an Tag 1 des ersten Zyklus durchgeführt (Baseline-Phase). Falls die körperlichen Untersuchungen länger als sieben Tage vor Tag 1 des ersten Zyklus stattfanden, wurden diese zur Baseline-Untersuchung wiederholt. Laboruntersuchungen sowie Schwangerschaftstests konnten auch bis zu 72 Stunden vor Randomisierung durchgeführt werden. Patienten, die die Baseline-Phase durchlaufen und die Einschlusskriterien erfüllt hatten, erhielten die Studienmedikation in der Randomisierungsphase. Die Randomisierungsphase begann zum Zeitpunkt der Randomisierung, endete zum Datenschnitt für die Analyse des primären Endpunkts (13.11.2016) und bestand aus einer Behandlungs- und Nachbeobachtungsphase. In der Behandlungsphase wurden die Patienten mittels 1:1 Randomisierung einer der zwei Studienmedikationen Lenvatinib oder Sorafenib zugeteilt. Im Lenvatinibarm erhielten Patienten mit einem Körpergewicht unter 60 kg täglich oral eine 8 mg Dosis (zwei 4 mg Kapseln) und ab 60 kg täglich oral eine 12 mg Dosis (drei 4 mg Kapseln) Lenvatinib. Im Sorafenibarm erhielten die Patienten zweimal täglich oral eine 400 mg Dosis (zwei 200 mg-Tabletten) Sorafenib (entsprechend einer Tagesgesamtdosis von 800 mg). In beiden Behandlungsarmen wurde die Studienmedikation im 28-Tage Zyklus bis zur „Off-Treatment“ Visite, die innerhalb von 30 Tagen nach der letzten Dosis der Studienmedikation durchgeführt wurde, verabreicht. Tumoruntersuchungen wurden innerhalb der 8. Woche nach Randomisierung und alle acht Wochen danach, oder wenn sie klinisch indiziert waren, bis zu einer Tumorprogression durchgeführt. Die Behandlung mit der Studienmedikation wurde abgebrochen, wenn es zu einer objektiv bewerteten Krankheitsprogression oder zur Entwicklung einer inakzeptablen Toxizität kam, auf Wunsch des Patienten oder nach Widerruf der Zustimmung des Patienten. Nach der „Off-Treatment“ Visite erfolgte die Nachbeobachtungsphase. In dieser Phase wurden bei Patienten, die weiter an der Studie teilnahmen und die keine Krankheitsprogression zeigten, alle acht Wochen Tumoruntersuchungen durchgeführt, bis es zur Progression oder zum Beginn einer anderen Krebstherapie kam. Wenn ein Patient die Studie vor der 8. Woche beendete, wurden die Tumoruntersuchungen so nah wie möglich zu Woche 8, aber vor dem Beginn einer weiteren Krebsbehandlung, durchgeführt. Außerdem wurden alle 12 Wochen das Überleben der Patienten sowie alle Krebsbehandlungen dokumentiert. Die Nachbeobachtungsphase endete mit der Beobachtung von insgesamt 700 Todesfällen in beiden Therapiearmen (prädefinierter Datenschnitt für den primären Endpunkt). Die abschließende Extensionsphase bestand ebenfalls aus einer Behandlungsphase und einer Nachbeobachtungsphase. Während der

Extensionsphase blieben solche Patienten, die zum Zeitpunkt des Datenschnitts für die Analyse des primären Endpunkts (700 Todesfälle; 13.11.2016) noch mit der Studienmedikation behandelt wurden, in dem bei der Randomisierung zugewiesenen Therapiearm. Patienten, die die Behandlung aufgrund einer bestätigten Progression oder aufgrund sonstiger gemäß Studienprotokoll vorgeschriebenen Gründe bereits während der Randomisierungsphase beendet hatten, gingen direkt in die Nachbeobachtungsphase über. Der Zeitraum der Nachbeobachtung begann unmittelbar nach Therapieende und dauerte bis zum Zeitpunkt des Todes oder bis zum Widerruf der Einwilligung des Patienten. Die Nachbeobachtung zum Langzeitüberleben wurde alle zwölf Wochen für alle Patienten durchgeführt. Patienten, die die Studie vor Krankheitsprogression beendet hatten, wurden alle acht Wochen bis zum Eintritt einer Progression oder dem Start einer anderen Krebsbehandlung nachbeobachtet (Eisai Inc. 2017b).

Randomisierung

Die Patienten wurden nach einem Randomisierungsverhältnis von 1:1 einem der zwei Studienarme zugeteilt (Arm A: Lenvatinib; Arm B: Sorafenib), stratifiziert nach „Region“ (zwei Ausprägungen: Asien-Pazifik, Westliche Regionen), „Makroskopische Invasion der Pfortader, Extrahepatische Ausbreitung oder beides“ (zwei Ausprägungen: Ja, Nein), „ECOG PS“ (zwei Ausprägungen: 1, 0) und „Körpergewicht“ (zwei Ausprägungen: <60 kg; ≥60 kg). Die Randomisierung wurde zentral mit Hilfe eines unabhängigen IxRS durchgeführt. Das Studienpersonal und die Patienten waren hinsichtlich der Studienbehandlung, infolge des Open-Label-Studiendesigns, nicht verblindet (Eisai Inc. 2017b).

Der Grund für das Open-Label Design der Studie 304 liegt in den spezifischen Leitlinien für das Toxizitätsmanagement von Lenvatinib und Sorafenib, die ein doppelblindes Studiendesign sehr schwierig gemacht hätten. Darüber hinaus gibt es Unterschiede in der Verabreichung der Wirkstoffe Lenvatinib (Kapseln) und Sorafenib (Tabletten) sowie in der Dosierung. Eine multiple Verabreichung von Placebo zu den bestimmten Zeitpunkten, um eine Verblindung aufrecht zu erhalten, wäre für die sehr schwer erkrankten Patienten verwirrend gewesen, was zu einem hohen Risiko einer Fehldosierung hätte führen können. Konsequenterweise wurde daher ein Open-Label Design gewählt, um die Sicherheit der Patienten und das Einhalten des Studienprotokolls zu gewährleisten (Eisai Inc. 2017b).

Patientenrelevante Endpunkte

Als primärer Endpunkt wurde in der Studie 304 das OS untersucht. Als patientenrelevante sekundäre Endpunkte wurden die Daten für das PFS, für die EORTC QLQ-C30 Symptom- und Funktionsskalen, für die Skalen des EORTC QLQ-HCC18 zur Messung der Morbidität und gesundheitsbezogener Lebensqualität und für die EQ-5D VAS erhoben. Als weitere patientenrelevante Endpunkte wurden Endpunkte zu Sicherheit und Verträglichkeit der Behandlung untersucht.

Die folgenden patientenrelevanten Endpunkte werden im vorliegenden Nutzendossier zur Ableitung des Zusatznutzen berichtet:

Mortalität:

- OS

Morbidität:

- PFS
- EORTC QLQ-C30 (Symptome)
    - Fatigue
    - Übelkeit und Erbrechen
    - Schmerz
    - Dyspnoe
    - Schlaflosigkeit
    - Appetitminderung
    - Obstipation
    - Diarrhö
- EORTC QLQ-HCC18 (Symptome)
    - Fatigue
    - Ikterus
    - Schmerz
    - Fieber
    - Abdominale Schwellung
- EQ-5D VAS

Gesundheitsbezogene Lebensqualität:

- EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität)
    - Physische Funktion
    - Rollenfunktionsniveau
    - Emotionale Funktion
    - Kognitive Funktion
    - Soziale Funktion
    - Globaler Gesundheitsstatus
    - Finanzielle Schwierigkeiten
- EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität)
    - Körperbild

- Ernährung
- Sexualleben

Sicherheit und Verträglichkeit:

- UE
- SUE
- UE, die zum Therapieabbruch führten
- UE mit CTCAE Grad ≥ 3

Patientencharakteristika

Die beiden Studienarme (Lenvatinib und Sorafenib) der Studie 304 waren vergleichbar bezüglich der demografischen Parameter und weiteren Charakteristika. Das Durchschnittsalter der Patienten betrug 61,3 Jahre im Lenvatinibarm und 61,2 Jahre im Sorafenibarm. Die Anzahl männlicher Patienten war ebenfalls vergleichbar und machte im Lenvatinibarm 84,7% und im Sorafenibarm 84,2% aller Patienten aus. Im Lenvatinibarm waren 28,2% der Patienten Weiße und 69,9% Asiaten. Im Vergleich dazu waren im Sorafenibarm 29,6% der Patienten Weiße und 68,5% Asiaten. Das durchschnittliche Gewicht und die durchschnittliche Größe betrugen im Lenvatinibarm 68,6 kg und 167,1 cm, im Sorafenibarm betrugen sie 68,1 kg und 167,2 cm.

Der Anteil der Patienten in Bezug auf den ECOG-PS, die New York Heart Association (NYHA)-Klassifikation und den Child-Pugh Score war in beiden Gruppen vergleichbar. Im Lenvatinibarm betrug der Anteil für den ECOG-PS mit dem Wert null und eins 63,6% resp. 36,4% und im Sorafenibarm 63,2% resp. 36,8%. Für die NYHA-Klassifikation lag der Anteil für Klasse I und II bei 7,7% resp. 1,7% im Lenvatinibarm. Im Sorafenibarm lagen die Anteile bei 9,2% resp. 1,3%. Der Anteil der Patienten mit einem Child-Pugh Score von 5, 6, 7 und 8 betrug im Lenvatinibarm 77,0%, 22,4%, 0,6% und 0,2%. Die entsprechenden Anteile lagen im Sorafenibarm bei 75,0%, 23,9%, 0,8% und 0,2%.

Auch bezüglich des Bestehens einer makroskopischen Invasion der Pfortader (engl. Macroscopic Portal Vein Invasion, MPVI), extrahepatischen Ausbreitung (engl. extrahepatic Spread, ES), MPVI oder ES oder beides oder Leberzirrhose sind beide Arme vergleichbar. Die Anteile im Lenvatinibarm lagen bei 22,8%, 60,9%, 68,8% und 50,8%. Im Sorafenibarm lagen die Anteile bei 18,9%, 62,0%, 70,6% und 48,5%.

Auch bezüglich der Krankheitsgeschichte sind beide Arme vergleichbar, lediglich für den im HCC prognostischen Faktor Alpha-Fetoprotein (AFP) und Patienten mit HCV als zugrundliegende Erkrankung besteht ein Ungleichgewicht. Der Anteil der Patienten mit einem Wert ≥ 200 ng/mL zu Baseline lag im Lenvatinibarm bei 46,4% und im Sorafenibarm bei 39,3%. Der Anteil der Patienten mit HCV als zugrundeliegende Erkrankung lag im Lenvatinibarm bei 19,0 % und im Sorafenibarm bei 26,5 % (Tabelle 4-8).

Patienten mit BCLC-Stadium B

Während des Beratungsgesprächs beim G-BA am 20.12.2017 wurde seitens der Geschäftsstelle darauf hingewiesen, dass hinsichtlich der Erstellung des Dossiers zur Nutzenbewertung darzulegen sei, aus welchen Gründen Patienten im BCLC-Stadium B in der pivotalen Studie nicht für eine TACE infrage kamen und inwieweit diese Gründe auf die deutsche Versorgungsrealität zutreffen. Diese Gründe werden im Folgenden diskutiert.

In die Studie 304 wurden neben Patienten im fortgeschrittenen Erkrankungsstadium BCLC-Stadium C, bei denen die Leitlinien eine systemische Therapie empfehlen (EASL 2018), auch Patienten im intermediären Stadium (BCLC-Stadium B) eingeschlossen (196/954 Patienten, 20,5 % der Studienpopulation). Diese Patienten sind gemäß aktueller Klassifizierung durch eine intrahepatische Erkrankung, die einem kurativem Verfahren (Resektion) nicht bzw. nicht mehr zugänglich ist sowie durch erhaltene Leberfunktion und fehlende Tumorsymptomatik bei Abwesenheit von makroskopischer vaskulärer Invasion und/oder extrahepatischer Streuung charakterisiert (EASL 2018). Der BCLC-Therapiealgorithmus sieht für die Patienten im BCLC-Stadium B als Standardbehandlung TACE vor (EASL 2018). Die Erkrankungen in diesem Stadium gelten allerdings als heterogen in Bezug auf Tumorlast und Leberfunktion (Sieghart et al. 2015), weswegen weitere Unterteilungen des BCLC-Stadium B zwecks Selektion der Patienten für die Behandlung vorgeschlagen wurden (Bolondi et al. 2012; Kudo et al. 2015). Eine TACE-Behandlung gilt am aussichtsreichsten bei Patienten mit solitärem Tumor bzw. Vorliegen einer kleinen Anzahl von Läsionen (EASL 2018), während bei Patienten mit bilobärem, multiplem Befall (≥ 7 Läsionen) bzw. bilobärem Befall mit großen Läsionen ein geringerer Nutzen zu erwarten ist (Kudo 2018). Des Weiteren nennen Leitlinien und Fachliteratur (EASL 2018; Raoul et al. 2011; Sieghart et al. 2015; Verslype et al. 2012) unterschiedlich abgestufte Kontraindikationen der TACE: In Bezug auf die Tumorerkrankung werden ausgedehnter (> 50 %) bzw. bilobärer Befall der Leber, große Knoten > 10 cm sowie maligne Portalvenenthrombose als absolute bzw. relative Ausschlusskriterien für die TACE genannt. Im Hinblick auf die Leberfunktion gelten dekompensierte Zirrhose (Child-Pugh B ≥8 inklusive Ikterus, hepatischer Enzephalopathie, Aszites und/oder hepatorenalem Syndrom), beeinträchtigter portalvenöser Blutfluss (Portalvenen-Thrombus, hepatofugaler Blutfluss) und ösophageale Varizen mit hohem Blutungsrisiko als Kontraindikationen. Weitere Ausschlusskriterien stellen Nierenfunktionseinschränkungen (Kreatinin ≥ 2 mg/dl bzw. Kreatinin-Clearance < 30 ml/min), nicht-behandelbare arteriovenöse Fisteln, schwere Komorbiditäten, die eine Einschränkung der Organfunktion zur Folge haben (z.B. kardiovaskuläre oder pulmonale Erkrankungen), sowie Gallengangokklusion und Inkompetenz der Vater-Papille dar. Letztlich ist die Anwendung des TACE-Verfahrens stets eine individuelle Entscheidung, die gemäß den Leitlinien interdisziplinär unter Einbezug eines Tumorboards getroffen werden soll (DGVS 2013; EASL 2018).

Ein Abbrechen der TACE-Behandlung wird nach zwei Anwendungen, in denen keine substantielle Tumornekrose erzielt wurde, empfohlen. Sofern in der Folgebehandlung einer progredienten Läsion keine deutliche Nekrose nach zuvor erfolgreicher Anwendung mehr erzielt werden kann, soll das Verfahren ebenfalls nicht wiederholt werden. Des Weiteren soll die Behandlung mit TACE bei Vorliegen einer nicht-behandelbaren Progression beendet

werden, worunter ein ausgedehnter Tumorbefall der Leber, extrahepatische Streuung oder vaskuläre Invasion, aber auch Beeinträchtigungen der Leberfunktion oder Verschlechterungen des Leistungsstatus bei weniger stark ausgeprägter Erkrankungsprogression zu verstehen sind. Auch die kumulativ verabreichte Dosis von Anthrazyklinen ist aufgrund der Kardiotoxizität zu berücksichtigen und sollte einen definierten Wert nicht überschreiten (EASL 2018). Die Nicht-Einsetzbarkeit von TACE in der klinischen Behandlungspraxis bei bestimmten Patientengruppen innerhalb des BCLC- Stadium B bzw. das Vorliegen einer Refraktärität gegenüber TACE wurde durch ein entsprechendes Einschlusskriterium („*not applicable for transarterial chemoembolization*“) in der Studie 304 berücksichtigt (Eisai Inc. 2017b). Die Beurteilung der Nicht-Eignung für TACE lag im Ermessen des Prüfarztes, wobei im Hinblick auf die o.e. Leitlinienempfehlung der Einbezug eines Tumorboards vorausgesetzt werden kann (DGVS 2013; EASL 2018). Eine detaillierte Überprüfung der Einzelpatientendaten aus der Baseline-Erhebung erlaubt es, diese Ermessensentscheidung in der überwiegenden Mehrzahl der Fälle objektiv nachzuvollziehen, wie im Folgenden dargestellt wird. In der Studie 304 wurden die Patienten im BCLC- Stadium B in vergleichbarem Ausmaß auf die Behandlungen mit Lenvatinib (104/478 Patienten, 21,8 %) und Sorafenib (92/476 Patienten, 19,3 %) randomisiert (Eisai Inc. 2018b). Mit einem Anteil von 80,1 % (157/196 Patienten) hatten sich die BCLC- Stadium B Patienten mehrheitlich bereits mindestens einer lokoregionären Vorbehandlung ihrer Tumorerkrankung unterzogen, wobei die transarterielle Embolisation, mit oder ohne Applikation eines Chemotherapeutikums, bei 84,7 % (133/157) der lokoregionär vorbehandelten Patienten eingesetzt worden war (Tabelle 4-9). Die verbleibenden 32,1 % (63/196) der BCLC- Stadium B Patienten ohne Vorbehandlung mit transarterieller Embolisation wiesen mindestens eine – in einigen Fällen mehrere – der folgenden absoluten und relativen, tumorbedingten Kontraindikationen eines transarteriellen Embolisationsverfahren auf (Eisai Inc. 2018b): 28 Patienten hatten eine hohe Tumorlast (16 Patienten mit einer Summe der Zielläsionsdurchmesser ≥ 10 cm und 12 Patienten mit einer Summe der Zielläsionsdurchmesser zwischen 7 cm bis < 10 cm und Vorliegen weiterer intrahepatischer Nichtzielläsionen), 19 Patienten wiesen Zielläsionen in mehr als einem Leberlappen auf, 14 Patienten wiesen Nichtzielläsionen in mehr als einem Leberlappen bzw. multiple Läsionen auf. Aufgetretene Kontraindikationen im Zusammenhang mit der Leberfunktion waren das Vorliegen von Aszites zur Baseline (2 Patienten), Enzephalopathie in der Anamnese (2 Patienten), Splenomegalie als Hinweis auf portale Hypertonie (4 Patienten), portale Hypertonie in der Krankengeschichte (4 Patienten) und ösophageale bzw. gastrische Varizen (14 Patienten). Aortenaneurysma, Myokardinfarkt, myokardiale Ischämie oder chronische obstruktive Lungenkrankheit stellten bedeutende Komorbiditäten bei 7 Patienten dar. Somit lässt sich zusammenfassend feststellen, dass die in Studie 304 eingeschlossenen BCLC- Stadium B Patienten in der ganz überwiegenden Mehrzahl (179 von 196 Patienten, 91,3 %) ein oder mehrere Ausschlusskriterien für eine TACE/TAE-Behandlung aufwiesen (Eisai Inc. 2018b): So hatten die Patienten überwiegend eine lokoregionale Vorbehandlung erhalten, wobei die Mehrzahl der Patienten mit einem transarteriellen Embolisationsverfahren behandelt wurde. Die Patienten, die nicht mit einem transarteriellen Embolisationsverfahren vorbehandelt waren, wiesen mehrheitlich einen massiven Tumorbefall bzw. Zeichen einer fortgeschrittenen Leberkrankung auf, die die Anwendung eines solchen Verfahrens ausschlossen.

Tabelle 4-9: Vorangegangene Tumorbehandlungen der Patienten im BCLC- Stadium B zu Baseline – Studie 304 mit dem zu bewertenden Arzneimittel

	Patienten im BCLC-Stadium B zu Baseline (N=196)
Patienten mit vorangegangener Tumorbehandlung, n (%)	157 (80,1)
Anzahl der vorangegangenen Behandlungen, n (%)	
1	68 (34,7)
2	47 (24,0)
3	28 (14,3)
4	12 (6,1)
≥5	2 (1,0)
Bezeichnung der vorangegangenen Behandlung[a]	
Intraarterielle Chemotherapie	21 (10,7)
Transarterielle (Chemo-) Embolisation	133 (67,9)
Radiofrequenzablation	67 (34,2)
Cryoablation	0 (0,0)
Perkutane Ethanolinjektion	14 (7,1)
Hepatektomie	52 (26,5)
Andere	17 (8,7)
Einsatz der vorangegangenen Behandlung[a], n (%)	
Neoadjuvant	8 (4,1)
Adjuvant	7 (3,6)
Therapeutisch	139 (70,9)
Konsolidierend	2 (1,0)
Erhaltend	3 (1,5)
Unbekannt	11 (5,6)

Die prozentualen Anteile beziehen sich auf die Gesamtzahl der Patienten im Stadium BCLC-B zu Baseline (ITT-Population).

a: Ein Patient kann bei mehreren Behandlungen gezählt werden.

BCLC: Barcelona Clinic Liver Cancer; n:Anzahl der Patienten mit einem Ereignis; N: Gesamtzahl der Patienten im Stadium BCLC-B zu Baseline (ITT-Population)

Quelle: (Eisai Inc. 2017b)

Übertragbarkeit auf den deutschen Versorgungskontext

Die Ergebnisse der Studie 304 sind aufgrund des Studiendesigns, der Studienpopulation und der Intervention auf den deutschen Versorgungskontext übertragbar.

Die Einschlusskriterien für die Studienpopulation entsprechen der Patientengruppe mit Leberzellkarzinom, die für eine systemische Therapie geeignet ist und im deutschen

Praxisalltag mit Lenvatinib gemäß Fachinformation behandelt wird. Wie in Modul 3 Abschnitt 3.2.1 beschrieben, liegt das mittlere Erkrankungsalter beim Leberzellkarzinom bei 69,9 Jahren für Männer und bei 72,1 Jahre für Frauen. Das durchschnittliche Alter der Studienpopulation der Studie 304 liegt etwas niedriger bei 61,3 Jahren. Dies ist ein übliches Phänomen, das in vielen onkologischen Studien auftritt. Es kann durch eine generell höhere Bereitschaft jüngerer Patienten an klinischen Studien teilzunehmen, erklärt werden. Zudem zeigt sich in den vorliegenden Subgruppenanalysen für den Endpunkt Gesamtüberleben keine Interaktion bezüglich des Alters. Bezüglich der unerwünschten Ereignisse finden sich ebenfalls keine relevanten Interaktionen bezüglich des Alters, so dass auch hier von einer guten Übertragbarkeit ausgegangen werden kann.

Des Weiteren entspricht die Dosierung und Darreichungsform von Lenvatinib in Studie 304 (einmal täglich oral 8 mg für Patienten mit einem Körpergewicht < 60 kg und 12 mg für Patienten mit einem Körpergewicht ≥ 60 kg) der im deutschen Praxisalltag gemäß Fachinformation für Lenvima® eingesetzten Dosierung und Darreichungsform (Eisai 2018a).

Zusammenfassend ist festzustellen, dass die Ergebnisse der Studie 304 auf den deutschen Versorgungskontext übertragbar sind. Hinweise auf Einschränkungen liegen nicht vor.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-10: Verzerrungspotenzial auf Studienebene – Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
304	ja	ja	nein	nein	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 entspricht nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie der Evidenzstufe Ib (G-BA 2018b). Das Verzerrungspotenzial wird als niedrig eingestuft.

Die Randomisierung wurde adäquat durchgeführt und die Gruppenzuteilung erfolgte rein zufällig mittels eines zentralen unabhängigen IxRS. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren oder – wie bei den Endpunkten zu Sicherheit und Verträglichkeit und den Subgruppenanalysen – gemäß Verfahrensordnung des G-BA berichtet wurden (z.B. die Berechnung der Effektmaße für den Vergleich der Behandlungsgruppen). Die Darstellung der Ergebnisse gibt keinen Anlass zur Vermutung, dass Ergebnisse selektiv berichtet wurden (Tabelle 4-10). Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschließt.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-11: Matrix der Endpunkte in der eingeschlossenen Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Morbidität	Gesundheits-bezogene Lebensqualität	Sicherheit und Verträglichkeit
	OS	PFS EORTC QLQ-C30 (Symptome) EORTC QLQ-HCC18 (Symptome) EQ-5D VAS	EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität)	UE[a]
304	ja	ja	ja	ja

a: berichtet als Anzahl Patienten mit UE, SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3

CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); SUE: Schwerwiegendes unerwünschtes Ereignis / schwerwiegende unerwünschte Ereignisse; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2017b

4.3.1.3.1 Endpunkt – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

– Ergebnisse der ITT-Analyse
– Zahl der Patienten, die in die Analyse eingegangen sind
– dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
– bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
– bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler
– Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

4.3.1.3.1.1 Endpunkt Mortalität – RCT

4.3.1.3.1.1.1 Gesamtüberleben (OS) – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-12: Operationalisierung des Endpunkts OS

Studie	Operationalisierung
304	OS wurde definiert als die Zeitspanne zwischen dem Zeitpunkt der Randomisierung bis zum Eintreten des Todes aufgrund jeglicher Ursache. Lost-to-follow-up-Patienten wurden zum letzten Zeitpunkt, an dem sie bekanntermaßen am Leben waren, zensiert. Überlebende Patienten wurden zum Zeitpunkt des Datenschnitts zensiert. Während des Studienverlaufs gab es zwei im Voraus geplante und im SAP definierte Zwischenanalysen. Die erste Zwischenanalyse wurde am 13. Mai 2015 durchgeführt, nachdem etwa 30% der notwendigen Fallzahl (entsprach 249 Todesfällen) in beiden Behandlungsarmen beobachtet wurde. Die zweite Zwischenanalyse wurde am 03. Februar 2016 durchgeführt, nachdem circa 70% der notwendigen Fallzahl (entsprach 521 Todesfällen) in beiden Behandlungsarmen beobachtet wurde. Basierend auf beiden von einem unabhängigen Statistikerteam durchgeführten Zwischenanalysen wurde vom DSMB empfohlen, die Studie ohne weitere Modifikationen fortzuführen. Die finale Analyse erfolgte am 13.November 2016. Die Analysen des Endpunkts OS wurden auf Basis der Daten der ITT-Population durchgeführt. <u>Hauptanalyse:</u> Für den Endpunkt OS wurde als primäre Analyse entsprechend dem SAP ein Nicht-Unterlegenheits-Test von Lenvatinib gegenüber Sorafenib durchgeführt. Dazu wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Die Efron-Methode wurde zur Korrektur von Bindungen verwendet. Für OS wurde zudem ein Überlegenheits-Test von Lenvatinib gegenüber Sorafenib durchgeführt. Dazu wurde ein zweiseitiger stratifizierter Log-Rank-Test, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 angewendet. Da die Studie nicht aufgrund einer der Zwischenanalysen gestoppt wurde, wurde die Nicht-Unterlegenheit bezüglich OS erstmals bei der finalen Analyse getestet. Nicht-Unterlegenheit galt als erfüllt, falls bei der finalen Analyse die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag. Falls Nicht-Unterlegenheit für den Endpunkt OS gezeigt wurde, sollte auf Überlegenheit getestet werden. Überlegenheit galt als erfüllt, wenn mithilfe des stratifizierten Log-Rank-Tests bei der finalen Analyse der zweiseitige p-Wert für den Test auf Gleichheit < 0,05 betrug. Das mediane OS wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das zum medianen OS gehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Überlebenswahrscheinlichkeiten wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt. <u>Sensitivitätsanalyse:</u> Für alle für das HCC als prognostisch identifizierten Patientencharakteristika wurden univariate Analysen mittels Cox-Proportional-Hazard-Modell durchgeführt und der Einfluss auf das HR analysiert. Zusätzlich wurde eine multivariate Analyse mittels Cox-Proportional-Hazard-Modell durchgeführt, wobei anhand einer vorwärts gerichteten schrittweisen Methode die Kovariaten gewählt wurden, die die Daten bezüglich des Wald-Tests, des Likelihood-Verhältnis-Tests und des Aikake’s Informationskriteriums am besten abbilden.

DSMB: Data Safety Monitoring Board; HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: interaktivem Voice and Web Response System; KI: Konfidenzintervall; OS: Overall Survival (Gesamtüberleben); SAP: statistischer Analyseplan;

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-13: Bewertung des Verzerrungspotenzials für den Endpunkt OS in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den patientenrelevanten Endpunkt OS beschrieben.

Das Verzerrungspotential wird als niedrig eingestuft, da die Beeinflussung hinsichtlich des Ergebnisses des Endpunkts OS auch bei einem Open-Label-Studiendesign durch das Wissen um die Behandlung weder durch den Patienten, noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts OS als niedrig eingestuft (Tabelle 4-13).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Ergebnisse der Hauptanalyse des Endpunkts OS für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt. Diese Ergebnisse werden im vorliegenden Dossier zur Ableitung des Zusatznutzens herangezogen.

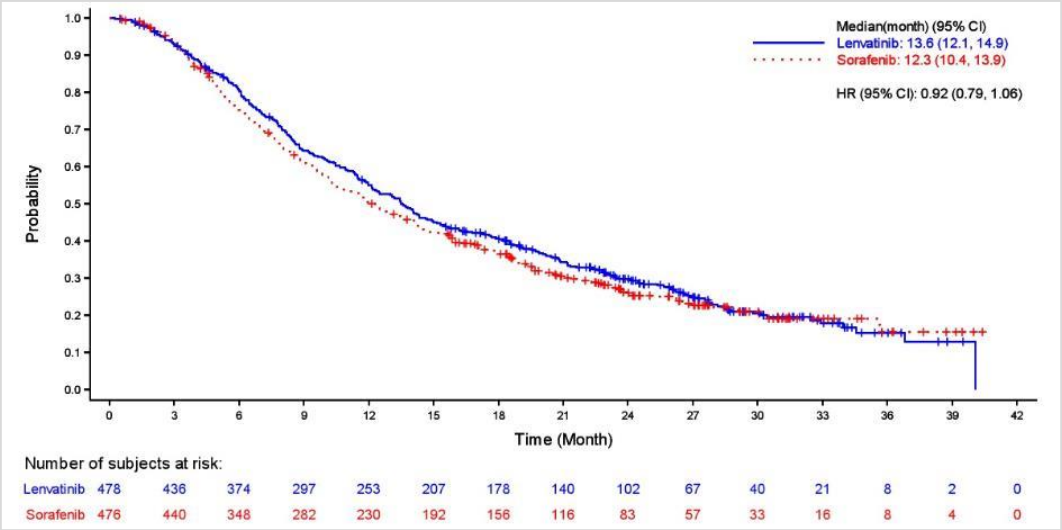


Abbildung 4-2: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; OS: Overall Survival (Gesamtüberleben)
Quelle: Eisai Inc. 2018c

Tabelle 4-14: Hauptanalyse: Ergebnisse für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel

Endpunkt	**Lenvatinib**		**Sorafenib**		**Lenvatinib vs. Sorafenib**
	n/N (%)	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
OS	351/478 (73,4)	13,6 [12,1; 14,9]	350/476 (73,5)	12,3 [10,4; 13,9]	0,92 [0,79; 1,06] 0,2526

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; OS: Overall Survival (Gesamtüberleben)

Quelle: Eisai Inc. 2018c

Die Kaplan-Meier-Kurve zeigt keine relevante Abweichung von der Proportionalitätsannahme (d.h. die Kurven laufen hinreichend parallel), welche die Grundlage des Cox-Proportional-Hazard-Modells für proportionale Hazards darstellt (Abbildung 4-2). Damit konnte die HR auf Basis des Cox-Proportional-Hazard-Modells berechnet werden. Auch die Kaplan-Meier-Schätzer für das mediane Überleben sind damit valide und interpretierbar.

Die Analysen des Endpunkts OS zeigten bezüglich des HR einen numerischen Vorteil für Lenvatinib gegenüber Sorafenib basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell (HR=0,92; 95%-KI [0,79; 1,06]; p=0,2526). Da die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag, konnte hiermit die Nicht-Unterlegenheit von Lenvatinib gegenüber Sorafenib belegt werden, was einen Überlebensvorteil von Lenvatinib gegenüber Placebo impliziert.

Die mediane Überlebensdauer unter Lenvatinib betrug 13,6 Monate (95%-KI [12,1; 14,9]), unter Sorafenib betrug sie 12,3 Monate (95%-KI [10,4; 13,9]) (Tabelle 4-14). Damit ergibt sich ein absoluter Vorteil im medianen Überleben von 1,3 Monaten zugunsten von Lenvatinib.

Die Studie 304 ist somit die erste positive, globale Phase-III-Studie in der Erstlinientherapie des HCC in den letzten 10 Jahren. Darüber hinaus ist die Studie 304 die erste positive Studie in HCC mit einer aktiven Kontrollgruppe. Seit der Zulassung von Sorafenib gab es vier Phase-III-Studien mit zielgerichteten Wirkstoffen, die keine Überlegenheit oder Nicht-Unterlegenheit in der Erstlinientherapie zeigen konnten (Kudo et al. 2018).

Insgesamt zeigen die Ergebnisse der Hauptanalyse eine Nicht-Unterlegenheit und einen numerischen Vorteil im Gesamtüberleben von Lenvatinib gegenüber Sorafenib.

Sensitivitätsanalysen

Tabelle 4-15: Sensitivitätsanalyse: Ergebnisse der adjustierten Analysen für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel

Adjustierung	**Lenvatinib vs. Sorafenib**
	HR[a] [95%-KI] **p-Wert**
OS	
Multivariate	0,81 [0,70; 0,94]
Alter (< 65, ≥ 65 bis < 75, ≥ 75)	0,92 [0,79; 1,07] 0,27
Geschlecht (männlich, weiblich)	0,92 [0,79; 1,06] 0,25
Region (Asien-Pazifik, Westliche Regionen)	0,92 [0,79; 1,06] 0,24
MPVI (ja, nein)	0,92 [0,79; 1,06] 0,25
ES (ja, nein)	0,91 [0,79; 1,06] 0,24
MPVI oder ES oder Beides (ja, nein)	0,91 [0,78; 1,05] 0,21
ECOG-PS (0, ≥ 1)	0,92 [0,80; 1,07] 0,29
Körpergewicht (< 60kg, ≥ 60kg)	0,92 [0,80;1,07] 0,30

Adjustierung	**Lenvatinib vs. Sorafenib**
	HR[a] [95%-KI] **p-Wert**
Baseline AFP (< 200 ng/mL, ≥ 200 ng/mL)	0,85 [0,73; 0,99] 0,04
Child-Pugh Score (5, ≥ 6)	0,93 [0,80; 1,08] 0,36
Betroffene Krankheitsstellen zu Baseline (1, 2, ≥3)	0,88 [0,76; 1,02] 0,09
Ätiologie (HBV, HCV, Alkohol)	0,91 [0,79; 1,06] 0,24
Zugrundeliegende Leberzirrhose (ja, nein)	0,92 [0,79; 1,06] 0,25
BCLC Staging (Stadium B oder C)	0,92 [0,79; 1,07] 0,27
Frühere Verfahren (ja, nein)	0,90 [0,77; 1,04] 0,15
Postprogressionstherapie	0,87 [0,75; 1,01] 0,06
Postprogressionsverfahren	0,89 [0,77; 1,04] 0,14
Postprogressionsmedikation	0,89 [0,76; 1,03] 0,11
a: Die HR mit dem dazugehörigen 95%-KI wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ und dem entsprechenden Baseline-Charakteristikum als Kovariate, berechnet; wenn ein Stratifizierungsfaktor als Kovariate verwendet wurde, wurde der entsprechende Stratifizierungsfaktor aus der stratifizierten Analyse ausgeschlossen AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; HR: Hazard Ratio; IxRS: Interaktives Voice and Web Response Systems; kg: Kilogramm; KI: Konfidenzintervall; mL: Milliliter; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); ng: Nanogramm; OS: Overall Survival (Gesamtüberleben) Quelle: Eisai Inc. 2018a	

Die Ergebnisse der Sensitivitätsanalysen, die auf uni- und multivariater Adjustierung basierten, zeigten konsistent gleichgerichtete Effekte wie die Ergebnisse der Hauptanalyse und bestätigen somit deren Robustheit. In Bezug auf die AFP adjustierte und multivariate adjustierte Analyse zeigte Lenvatinib im Vergleich zu Sorafenib sogar eine statistisch signifikante Überlegenheit im OS. Diese Heterogenität der Ergebnisse der Hauptanalyse und der adjustierten Analysen kann auf das Ungleichgewicht in den Behandlungsarmen bezüglich des für das HCC prognostischen Faktors AFP zurückgeführt werden. Der Anteil der Patienten mit einem Wert ≥ 200 ng/mL lag im Lenvatinibarm bei 46,4% und im Sorafenibarm bei 39,3%.

Unter Berücksichtigung der dargestellten Sensitivitätsanalysen zeigen die Ergebnisse zum OS nicht nur eine Nicht-Unterlegenheit und einen numerischen Vorteil im Gesamtüberleben

sondern auch eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib bezüglich der AFP adjustierten und der multivariaten adjustierten Analyse.

Postprogressionstherapie

Tabelle 4-16: Postprogressionstherapie: Ergebnisse für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Kategorie	**n/N (%)**	**Median[a] [95% KI]**	**n/N (%)**	**Median[a] [95% KI]**	**HR[b] [95% KI]**
OS					
Alle	351/478 (73,4)	13,6 [12,1; 14,9]	350/476 (73,5)	12,3 [10,4; 13,9]	0,92 [0,79; 1,06]
Postprogressionstherapie					
Patienten mit Postprogressionstherapie	143/206 (69,4)	19,5 [15,7; 23,0]	175/243 (72,0)	17,0 [14,2; 18,8]	0,84 [0,67; 1,06]
Patienten ohne Postprogressionstherapie	208/272 (76,5)	10,5 [8,6; 12,2]	175/233 (75,1)	7,9 [6,6; 9,7]	0,91 [0,74; 1,11]
Anteil Patienten mit Postprogressionstherapie					
	n/N (%)		n/N (%)		
	206/478 (43,1)		243/476 (51,1)		

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; OS: Overall Survival (Gesamtüberleben)

Quelle: Eisai Inc. 2017b.

In der Studie 304 war in beiden Behandlungsarmen das mediane OS 9 Monate länger für Patienten, die eine Postprogressionstherapie erhielten als für Patienten ohne Postprogressionstherapie. Unter der Behandlung mit Lenvatinib betrug das mediane OS 19,5 Monate (95%-KI [15,7; 23,0]) für Patienten mit Postprogressionstherapie und 10,5 Monate (95%-KI [8,6; 12,2]) für Patienten ohne Postprogressionstherapie. Im Sorafenibarm lag das mediane OS bei 17,0 Monaten (95%-KI [14,2; 18,8]) für Patienten mit und bei 7,9 Monaten (95%-KI [6,6; 9,7]) für Patienten ohne Postprogressionstherapie. Das OS für beide Teilpopulationen (mit und ohne Postprogressionstherapie) war also circa 2,5 Monate länger unter Lenvatinib im Vergleich zu Sorafenib (Eisai Inc. 2017b).

Der häufigere Einsatz von Postprogressionstherapien im Sorafenibarm könnte für das längere OS unter der Behandlung mit Sorafenib in der Studie 304 im Vergleich zu historischen Sorafenib Studien verantwortlich sein. Bei Anpassung an die Postprogressionstherapien (eine im SAP vordefinierte Analyse), liegt die HR für den Endpunkt OS bei 0,87 (95%-KI [0,75; 1,01]) zugunsten von Lenvatinib (Eisai Inc. 2018a).

In der Studie 304 gab es ein Ungleichgewicht zwischen den beiden Behandlungsarmen im Anteil der Patienten mit einer Postprogressionstherapie in der Nachbeobachtungsphase. Im Sorafenibarm lag der Anteil bei 51,1%, im Lenvatinibarm lag er bei 43,1%. Darüber hinaus war der Anteil an Patienten, die eine Postprogressionstherapie mit Prüfpräparaten wie Regorafenib erhielten unter der Behandlung mit Sorafenib mit 9,5% deutlich höher als unter der Behandlung mit Lenvatinib (3,1%) (Eisai Inc. 2017b). In den Zweitlinienstudien mit Prüfpräparaten war eine Sorafenib-Vortherapie oftmals ein Einschlusskriterium (siehe z.B. die Regorafenib RESORCE Studie (Bruix et al. 2017) und die Cabozantinib CELESTIAL Studie (Abou-Alfa et al. 2015)). Am Ende der Behandlungsphase waren die Patienten im Sorafenibarm der Studie 304 somit geeignet für diese Zweitlinienstudien. Hingegen waren Patienten aus dem Lenvatinibarm der Studie 304 für Studien, die eine Sorafenib-Vortherapie erforderten, nicht geeignet. Folglich ist das beobachtete Ungleichgewicht in der Behandlung mit neuen Wirkstoffen ein Artefakt des klinischen Studiendesigns und kein Abbild der klinischen Praxis.

Die häufigste medikamentöse Postprogressionstherapie im Lenvatinibarm war eine Behandlung mit Sorafenib (77,6%) (Eisai Inc. 2017b).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2 Endpunkt Morbidität – RCT

4.3.1.3.1.2.1 Progressionsfreies Überleben (PFS) – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-17: Operationalisierung des Endpunkts PFS

Studie	Operationalisierung
304	PFS ist definiert als Zeitspanne zwischen dem Zeitpunkt der Randomisierung bis zum Zeitpunkt einer dokumentierten Krankheitsprogression oder bis zum Zeitpunkt des Todes aufgrund jeglicher Ursache, abhängig davon, welches dieser beiden Ereignisse zuerst eintrat. Der Zeitpunkt der objektiven Krankheitsprogression war definiert als der früheste Zeitpunkt einer radiologisch bestätigten Krankheitsprogression. Patienten, die die Therapie aufgrund einer Krankheitsprogression beendeten oder bei denen keine radiologische Bestätigung für die Krankheitsprogression vorlag, wurden zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert. Patienten, die eine neue Therapie anfingen, oder die nach mehr als einer verpassten Visite oder nach bis zu 28 Tagen der letzten Dosis der Studienmedikation verstarben oder bei denen eine Progression auftrat, wurden ebenfalls zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert. Die Analysen des Endpunkts PFS wurden auf Basis der Daten der ITT-Population durchgeführt. <u>Hauptanalyse:</u> Für den Endpunkt PFS wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet. Der Unterschied im PFS zwischen Lenvatinib und Sorafenib wurde mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests analysiert, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert. Das mediane PFS wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das zum medianen PFS gehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für ein progressionsfreies Überleben wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt. Die Beurteilung der Krankheitsprogression erfolgte anhand der mRECIST-Kriterien durch IIR. Die Auswertung des IIR stellt auf Grund ihres verblindeten und unabhängigen Charakters im Folgenden die Hauptanalyse dar. <u>Ergänzende Analyse:</u> Ergänzend werden Analysen zur ORR dargestellt, um die Ergebnisse zum PFS zu untermauern. Dazu wurde die ORR (basierend auf den mRECIST IIR Daten) für beide Behandlungsarme bestimmt und der Unterschied zwischen Lenvatinib und Sorafenib mittels Odds Ratio (OR), dem relativen Risiko (engl. Risk Ratio, RR) und der Risikodifferenz (engl. Risk Rifference, RD) und den dazugehörigen 95%-KIs analysiert. Die OR und die statistische Signifikanz des Unterschieds in der ORR zwischen Lenvatinib und Sorafenib wurde mittels stratifiziertem Cochran-Mantel-Haenszel Chi-Quadrat-Test, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 untersucht. Bei der Analyse der ORR handelt es sich um eine ergänzende Analyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden. <u>Sensitivitätsanalyse</u> Die Auswertungen durch den Prüfarzt anhand der mRECIST-Kriterien und durch IIR anhand der RECIST v.1.1-Kriterien werden zusätzlich als Sensitivitätsanalysen für OS und PFS dargestellt. Die Ergebnisse der Sensitivitätsanalysen werden nicht zur Ableitung des Zusatznutzens herangezogen.

HR: Hazard Ratio; ITT: Intention-to-treat; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; OR: Odds Ratio; ORR: Objective Response Rate (Objektive Ansprechrate); PFS: Progression-Free Survival (progressionsfreies Überleben); RD: Risikodifferenz; RECIST: Response Evaluation Criteria in Solid Tumors; RR: Risk Ratio

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-18: Bewertung des Verzerrungspotenzials für den Endpunkt PFS in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den Endpunkt PFS beschrieben.

Das Verzerrungspotential wird als niedrig eingestuft, da der Endpunkt PFS mit klar definierten, objektiven Kriterien erfasst wurde (mRECIST). Wegen der Bewertung der Progression durch ein verblindetes unabhängiges Reviewboard ist keine Beeinflussung aufgrund des Open-Label-Studiendesigns zu erwarten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschließt. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts PFS (IIR Bewertung) als niedrig eingestuft (Tabelle 4-18).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Ergebnisse der Haupt- und Sensitivitätsanalyse sowie der ergänzenden Analyse des Endpunkts PFS für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt.

Hauptanalyse

Im Folgenden werden die Ergebnisse für den patientenrelevanten Endpunkt PFS auf Basis des IIR per mRECIST dargestellt. Diese Ergebnisse werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

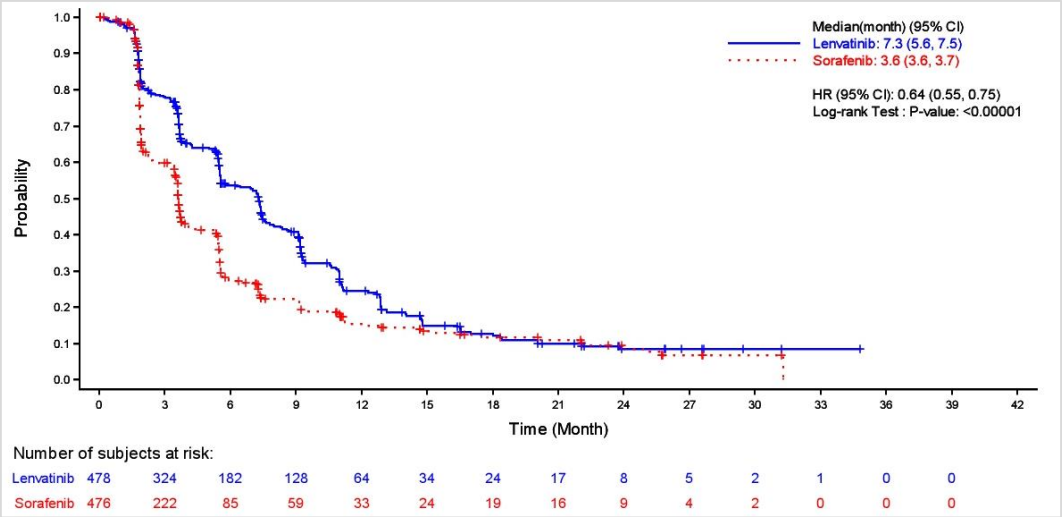


Abbildung 4-3: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); mRECIST: modified Response Evaluation Criteria in Solid Tumors; PFS: Progression-Free Survival (progressionsfreies Überleben)

Quelle: Eisai Inc. 2018c

Tabelle 4-19: Hauptanalyse: Ergebnisse für den Endpunkt PFS (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
PFS	311/478 (65,1)	7,3 [5,6; 7,5]	323/476 (67,9)	3,6 [3,6; 3,7]	0,64 [0,55; 0,75] <0,0001

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt

b: Die HR mit dem dazugehörigen 95%-KI wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

HR: Hazard Ratio; ITT: Intention-to-treat; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; n: Anzahl der Patienten mit einem Ereignis;

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
N: Anzahl der Patienten in der ITT-Population; PFS: Progression-Free Survival (progressionsfreies Überleben) Quelle: Eisai Inc. 2018c					

Die Hauptanalyse des Endpunkts PFS auf Basis des IIR per mRECIST zeigte im HR eine statistisch signifikante Überlegenheit bezüglich Risikoreduktion einer Progression von Lenvatinib gegenüber Sorafenib basierend auf dem Cox-Proportional-Hazard-Modell (HR 0,64; 95%-KI [0,55; 0,75]; p<0,0001) (Tabelle 4-19). Unter der Behandlung mit Lenvatinib wurde somit eine 36%-ige Verringerung des Progressionsrisikos im Vergleich zur Behandlung mit Sorafenib erreicht.

In Bezug auf die mediane Zeit bis zur Krankheitsprogression war das mediane PFS im Lenvatinibarm mit 7,3 Monaten (95%-KI [5,6; 7,5]) mehr als doppelt so lang wie im Sorafenibarm (3,6 Monaten (95%-KI [3,6; 3,7])) (Tabelle 4-19). Somit ergab sich ein klinisch relevanter absoluter Vorteil von 3,7 Monaten für das mediane PFS zugunsten von Lenvatinib.

Insgesamt belegen die Ergebnisse zum PFS eine statistisch signifikante und klinisch relevante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Verzögerung des Voranschreitens der Erkrankung.

Sensitivitätsanalysen

Im Folgenden werden die Ergebnisse für den Endpunkt PFS auf Basis des IIR per RECIST v.1.1 und auf Basis der Bewertung durch den Prüfarzt per mRECIST zur Unterstützung der Hauptanalyse des PFS dargestellt. Diese werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

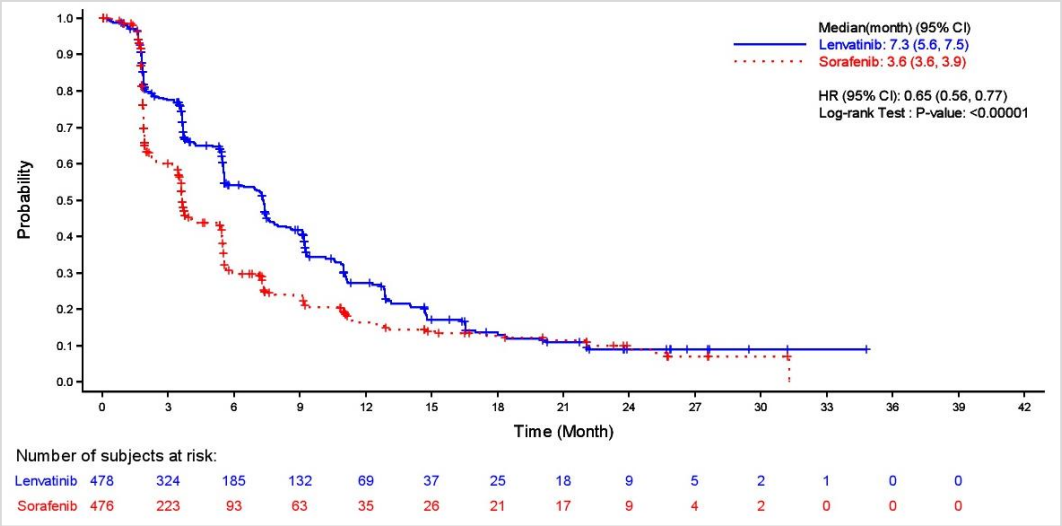


Abbildung 4-4: Sensitivitätsanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (IIR, RECIST v.1.1) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); PFS: Progression-Free Survival (progressionsfreies Überleben); RECIST: Response Evaluation Criteria in Solid Tumors

Quelle: Eisai Inc. 2018c

Tabelle 4-20: Sensitivitätsanalyse: Ergebnisse für den Endpunkt PFS (IIR, RECIST v.1.1) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
PFS	307/478 (64,2)	7,3 [5,6; 7,5]	320/476 (67,2)	3,6 [3,6; 3,9]	0,65 [0,56; 0,77] <0,0001

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

HR: Hazard Ratio; ITT: Intention-to-treat; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; PFS: Progression-Free Survival (progressionsfreies Überleben); RECIST: Response Evaluation Criteria in Solid Tumors

Quelle: Eisai Inc. 2018c

Das Ergebnis der Sensitivitätsanalyse des Endpunkts PFS auf Basis des IIR per RECIST v.1.1 zeigte sich konsistent zum Ergebnis der Hauptanalyse des Endpunkts PFS mittels IIR per mRECIST. Auch hier ergab sich ein Effektschätzer zugunsten von Lenvatinib. Lenvatinib zeigte im Vergleich zu Sorafenib bezogen auf das Risiko einer Progression basierend auf dem Cox-Proportional-Hazard-Modell (HR 0,65; 95%-KI [0,56; 0,77]; p<0,0001) eine statistisch signifikante und klinisch relevante Überlegenheit.Bezogen auf das mediane PFS (Lenvatinib: 7,3 Monate; 95%-KI [5,6; 7,5]; Sorafenib: 3,6 Monate; 95%-KI [3,6; 3,9]) ergab sich für Lenvatinib einabsoluter Vorteil von 3,7 Monaten. Die Sensitivitätsanalyse stützt somit die Robustheit der Hauptanalyse (Tabelle 4-20).

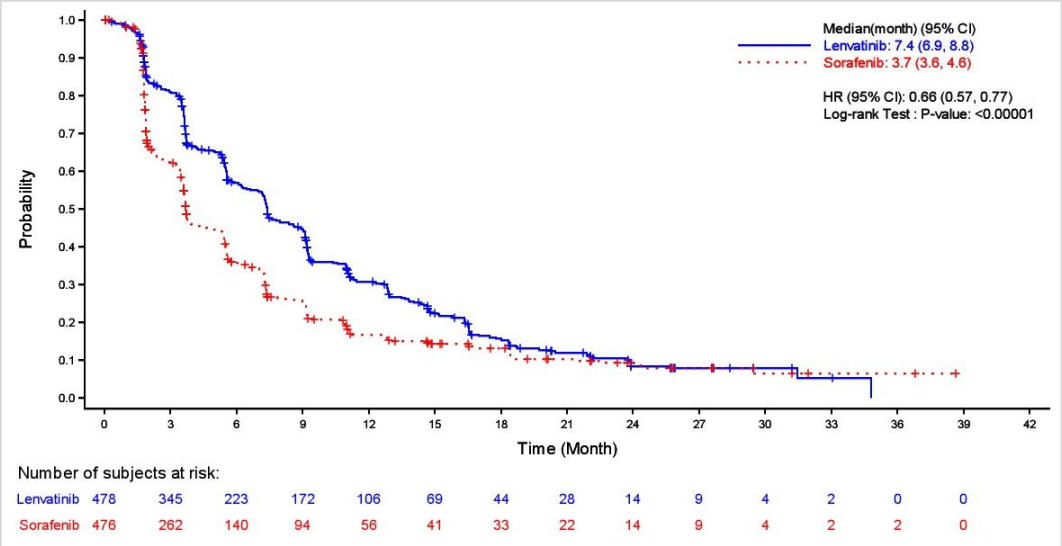


Abbildung 4-5: Sensitivitätsanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (Prüfarzt, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; mRECIST: modified Response Evaluation Criteria in Solid Tumors; PFS: Progression-Free Survival (progressionsfreies Überleben)
Quelle: Eisai Inc. 2018c

Tabelle 4-21: Sensitivitätsanalyse: Ergebnisse für den Endpunkt PFS (Prüfarzt, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
PFS	349/478 (73,0)	7,4 [6,9; 8,8]	367/476 (77,1)	3,7 [3,6; 4,6]	0,66 [0,57; 0,77] <0,0001
a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt					

	Lenvatinib		Sorafenib		Lenvatinib vs. Sorafenib
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
b: Die HR mit dem dazugehörigen 95%-KI wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; PFS: Progression-Free Survival (progressionsfreies Überleben) Quelle: Eisai Inc. 2018c					

Das Ergebnis der Sensitivitätsanalyse des Endpunkts PFS basierend auf der Tumorbewertung durch den Prüfarzt per mRECIST zeigte sich ebenfalls konsistent zum Ergebnis der Hauptanalyse. Auch hier zeigte die Therapie mit Lenvatinib im Vergleich zu Sorafenib bezogen auf das Risiko einer Progression basierend auf dem Cox-Proportional-Hazard-Modell (HR 0,66; 95%-KI [0,57; 0,77]; p<0,0001) eine statistisch signifikante und klinisch relevante Überlegenheit. Bezogen auf das mediane PFS (Lenvatinib: 7,4 Monate; 95%-KI [6,9; 8,8]; Sorafenib: 3,7 Monate; 95%-KI [3,6; 4,6]) ergab sich für Lenvatinib ein absoluter Vorteil von 3,7 Monaten. Auch diese Sensitivitätsanalyse stützt somit die Robustheit der Hauptanalyse (Tabelle 4-21).

Die Sensitivitätsanalysen zur weiteren Absicherung der Ergebnisse wiesen jeweils eine HR auf, die eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib belegt. Damit bestätigen die Ergebnisse der Sensitivitätsanalysen die Robustheit der Ergebnisse der Hauptanalyse.

Ergänzende Analyse

Im Folgenden werden die Ergebnisse für den Endpunkt ORR auf Basis der IIR per mRECIST und per RECIST v.1.1 sowie basierend auf der Tumorbewertung durch den Prüfarzt per mRECIST dargestellt, um die Ergebnisse des Endpunkts PFS zu untermauern. Diese werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

Tabelle 4-22: Ergänzende Analyse: Ergebnisse für den Endpunkt ORR (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib	**Sorafenib**	**Lenvatinib vs. Sorafenib**			
Endpunkt	**n/N (%)**	**n/N (%)**	**RR [95% KI]**	**OR [95% KI]**	**RD [95% KI]**	**p-Wert[a]**
ORR	194/478 (40,6)	59/476 (12,4)	3,27 [2,52; 4,26]	5,01 [3,59; 7,01]	28,2 [22,9; 33,5]	<0,0001

a: Der p-Wert bezieht sich auf den Effektschätzer OR und wird anhand des Cochrane-Mantel-Haenszel-Tests ermittelt, stratifiziert nach den IxRS Stratifizierungsfaktoren.

IIR: Independent Imaging Review (unabhängige radiologische Bewertung); IxRS: Interaktives Voice and Web Response Systems; ITT: Intention-to-treat; KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; OR: Odds Ratio; ORR: Objective Response Rate (Objektive Ansprechrate); RD: Risikodifferenz; RR: Relatives Risiko

Quelle: Eisai Inc. 2018c

Die Analyse des Endpunkts ORR basierend auf dem IIR per mRECIST zeigte unter Lenvatinib mit 40,6% eine deutlich höhere Ansprechrate als unter Sorafenib mit 12,5%. Dabei ergab sich aus einer OR von 5,01 (95%-KI [3,59; 7,01]) ein signifikanter p-Wert kleiner 0,0001. Bezüglich des Endpunkts ORR zeigte Lenvatinib somit eine statistisch signifikante Überlegenheit gegenüber Sorafenib (Tabelle 4-22).

Tabelle 4-23: Sensitivitätsanalyse: Ergebnisse für den Endpunkt ORR (IIR, RECIST v.1.1) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib	**Sorafenib**	**Lenvatinib vs. Sorafenib**			
Endpunkt	**n/N (%)**	**n/N (%)**	**RR [95% KI]**	**OR [95% KI]**	**RD [95% KI]**	**p-Wert[a]**
ORR	90/478 (18,8)	31/476 (6,5)	2,89 [1,96; 4,26]	3,34 [2,17; 5,14]	12,3 [8,2; 16,5]	<0,0001

a: Der p-Wert bezieht sich auf den Effektschätzer OR und wird anhand des Cochrane-Mantel-Haenszel-Tests ermittelt, stratifiziert nach den IxRS Stratifizierungsfaktoren.

IIR: Independent Imaging Review (unabhängige radiologische Bewertung); IxRS: Interaktives Voice and Web Response Systems; ITT: Intention-to-treat; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; OR: Odds Ratio; ORR: Objective Response Rate (Objektive Ansprechrate); RD: Risikodifferenz; RECIST: Response Evaluation Criteria in Solid Tumors; RR: Relatives Risiko

Quelle: Eisai Inc. 2018c

Auch bei der Analyse des Endpunkts ORR basierend auf dem IIR per RECIST v.1.1 lag die Ansprechrate von Lenvatinib mit 18,8% deutlich höher als die von Sorafenib mit 6,5%. Die sich daraus ergebene OR zeigte mit 3,34 (95%-KI [2,17; 5,14]; p<0,0001) eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die Analyse unter Verwendung der RECIST v.1.1. Kriterien zeigte in beiden Behandlungsarmen eine geringere Ansprechrate als bei der Verwendung der mRECIST Kriterien. Diese geringere Ansprechrate

kann damit begründet werden, dass die Tumorreduktion beim HCC durch die RECIST v.1.1 Kriterien nicht adäquat gemessen wird, wie im Abschnitt 4.2.5.2 dargestellt. Allerdings bestätigt auch die Analyse per RECIST v.1.1. das deutlich und statistisch signifikant erhöhte Ansprechen unter der Behandlung mit Lenvatinib im Vergleich zur Behandlung mit Sorafenib. Das Ergebnis der RECIST v.1.1. IIR Analyse ist somit konsistent zum Ergebnis der mRECIST IIR Analyse und stützt damit deren Robustheit (Tabelle 4-23).

Tabelle 4-24: Sensitivitätsanalyse: Ergebnisse für den Endpunkt ORR (Prüfarzt, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib	**Sorafenib**	**Lenvatinib vs. Sorafenib**			
Endpunkt	**n/N (%)**	**n/N (%)**	**RR [95% KI]**	**OR [95% KI]**	**RD [95% KI]**	**p-Wert[a]**
ORR	115/478 (24,1)	44/476 (9,2)	2,60 [1,88; 3,60]	3,13 [2,15; 4,56]	14,8 [10,2; 19,4]	<0,0001

a: Der p-Wert bezieht sich auf den Effektschätzer OR und wird anhand des Cochrane-Mantel-Haenszel-Tests ermittelt, stratifiziert nach den IxRS Stratifizierungsfaktoren.

ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; OR: Odds Ratio; ORR: Objective Response Rate (Objektive Ansprechrate); RD: Risikodifferenz; RR: Relatives Risiko

Quelle: Eisai Inc. 2018c

Auch das Ergebnis der Analyse basierend auf der Tumorbewertung durch den Prüfarzt per mRECIST zeigte sich konsistent zum Ergebnis der unabhängigen radiologischen Auswertung per mRECIST (Tabelle 4-22). Lenvatinib zeigte eine deutlich höhere Ansprechrate (24,1%) als Sorafenib (9,4%), wobei die OR mit 3,13 (95%-KI [2,15; 4,56]; p<0,0001) ebenfalls eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib belegt (Tabelle 4-24).

Tabelle 4-25: Ergebnisse für die Endpunkte CR und PR aus Studie 304 mit dem zu bewertenden Arzneimittel

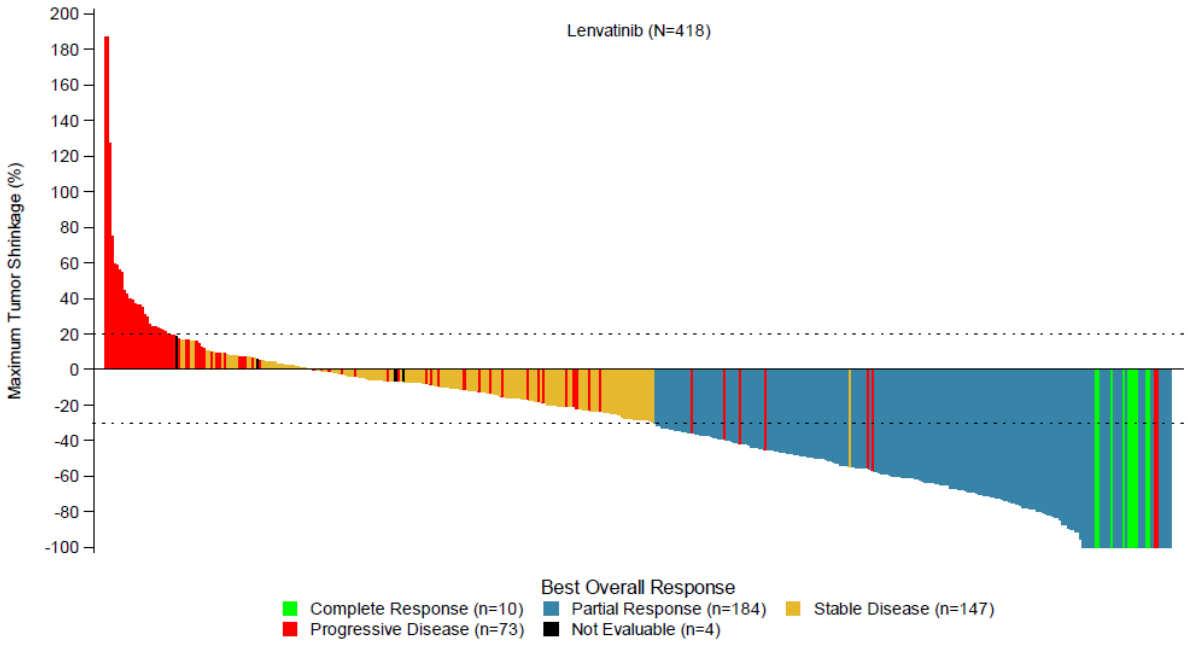
Endpunkte	**IIR**				**Prüfarzt**	
	mRECIST		**RECIST v1.1**		**mRECIST**	
	Lenvatinib (N=478) n(%)	**Sorafenib (N=476) n(%)**	**Lenvatinib (N=478) n(%)**	**Sorafenib (N=476) n(%)**	**Lenvatinib (N=478) n(%)**	**Sorafenib (N=478) n(%)**
CR	10 (2,1)	4 (0,8)	2 (0,4)	1 (0,2)	6 (1,3)	2 (0,4)
PR	184 (38,5)	55 (11,6)	88 (12,4)	30 (6,3)	109 (22,8)	42 (8,8)

CR: Complete Response (vollständiges Ansprechen); IIR: Independent Imaging Review (unabhängige radiologische Bewertung) ; mRECIST: modified Response Evaluation Criteria in Solid Tumors; ITT: Intention-to-treat; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; PR: Partial Response (partielles Ansprechen); RECIST: Response Evaluation Criteria in Solid Tumors

Quelle: Eisai Inc. 2017a

Auch die Betrachtung der Patienten, die unter der Behandlung eine CR (vollständiges Ansprechen) bzw. eine PR (partielles Ansprechen) zeigten, bestätigten den Vorteil von Lenvatinib gegenüber Sorafenib (Tabelle 4-25). Unter der Behandlung mit Lenvatinib zeigten 10 Patienten eine vollständige Remission (per IIR mRECIST), d.h. ein vollständiges Verschwinden der Läsionen. Unter der Behandlung mit Sorafenib wurde die vollständige Remission bei 4 Patienten beobachtet.

Das bessere Ansprechen des Tumors unter der Behandlung mit Lenvatinib im Vergleich zur Behandlung mit Sorafenib ist auch in den in Abbildung 4-6 gezeigten Waterfall-Plots sichtbar, die die prozentuale Veränderung der aufsummierten Durchmesser der Zielläsionen für die einzelnen Patienten darstellt (Eisai Inc. 2017b).



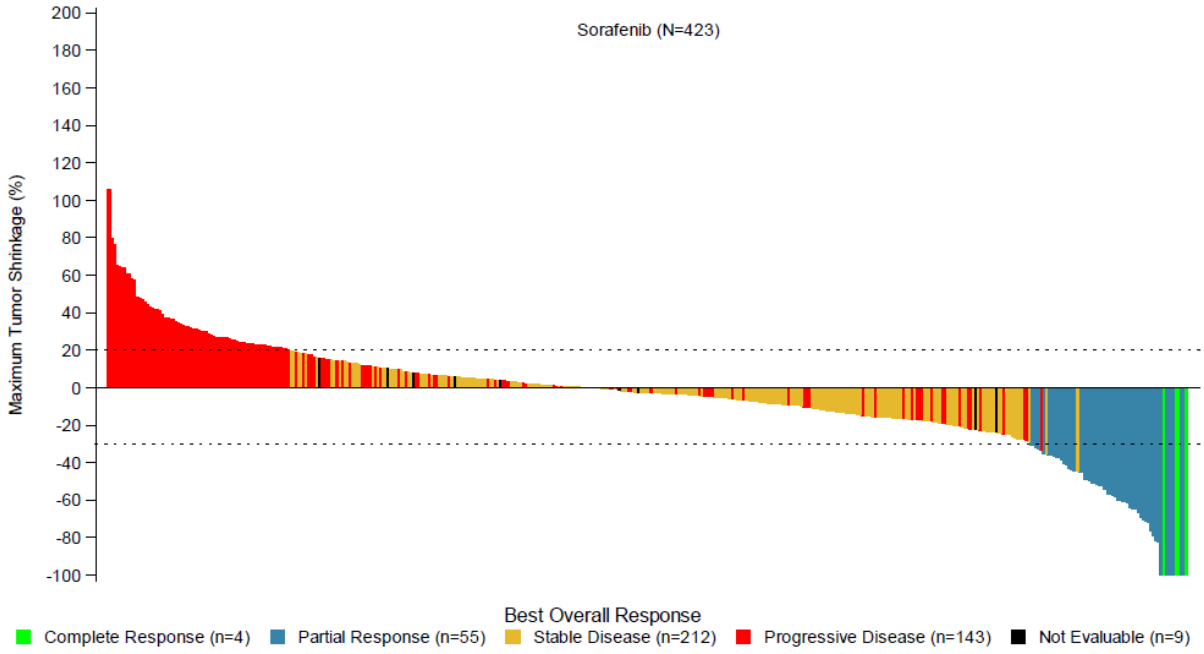


Abbildung 4-6: Waterfall Plots für die prozentuale Veränderung der aufsummierten Durchmesser der Zielläsionen für die einzelnen Patienten (IIR, mRECIST)

ITT: Intention-to-treat; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); mRECIST: modified Response Evaluation Criteria in Solid Tumors; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population;
Quelle: Eisai Inc. 2018c

Alle drei im Rahmen der Studie 304 durchgeführten Auswertungen des Endpunkts ORR zeigten eine klare und statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die Ergebnisse zur Ansprechrate bestätigen die Robustheit der Ergebnisse zum Endpunkt PFS.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2.2 EORTC QLQ-C30 (Symptome) – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-26: Operationalisierung des Endpunkts EORTC QLQ-C30 (Symptome)

Studie	Operationalisierung
304	Der Fragebogen EORTC QLQ-C30 (Symptome) ist in der Onkologie ein häufig eingesetztes Instrument zur Bewertung der Morbidität. Die gestellten Fragen beziehen sich dabei auf die Symptome Fatigue, Übelkeit und Erbrechen, Schmerz, Dyspnoe, Schlaflosigkeit Appetitminderung, Obstipation und Diarrhö. Die einzelnen Items werden dabei auf einer Skala von 1 bis 4 bewertet, die zur Auswertung in eine Skala von 0 bis 100 transformiert wird. Höhere Werte auf den Symptomskalen bedeuten eine Verschlechterung der Symptome (Giesinger et al. 2016). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen (Cocks et al. 2011; Osoba et al. 1998). Die Analysen des Endpunkts EORTC QLQ-C30 (Symptome) wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. <u>Hauptanalyse:</u> Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt. <u>Sensitivitätsanalyse:</u> Die Robustheit der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms, gemessen auf den Skalen des EORTC QLQ-C30 Fragebogens, wird mit Hilfe der mittleren Veränderung zu Baseline im Score untermauert. Dazu wurden der Mittelwert zu Baseline und der Least Square(LS)-Mittelwert der Änderung zu Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zu Baseline analysiert. Bei der Analyse der mittleren Veränderung zu Baseline im Score des jeweiligen Symptoms bzw. der jeweiligen Funktion handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; LS: Least Square

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-27: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-C30 (Symptome) in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-C30 beschrieben.

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95% (Tabelle 4-28). Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-C30 als hoch eingestuft (Tabelle 4-27).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Rücklaufquoten und die Ergebnisse der Haupt- und Sensitivitätsanalyse des Fragebogens EORTC QLQ-C30 (Symptome) für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt.

Rücklaufquoten

Tabelle 4-28: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**n (%)**	**N**	**n (%)**
Fatigue	478	446 (93,3)	476	451 (94,7)
Übelkeit und Erbrechen	478	446 (93,3)	476	451 (94,7)
Schmerz	478	446 (93,3)	476	451 (94,7)
Dyspnoe	478	444 (92,9)	476	450 (94,5)
Schlaflosigkeit	478	446 (93,3)	476	451 (94,7)
Appetitminderung	478	446 (93,3)	476	451 (94,7)
Obstipation	478	445 (93,1)	476	450 (94,5)
Diarrhö	478	444 (92,9)	476	450 (94,5)
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ITT: Intention-to-treat; n: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert; N: Anzahl der Patienten in der ITT-Population Quelle: SCO:SSiS 2018				

Die Rücklaufquoten des Fragebogens EORTC QLQ-C30 (Symptome) sind in Tabelle 4-28 dargestellt. Hierbei handelt es sich um den Anteil derjenigen Patienten in der ITT-Population mit einem Baseline-Wert und mindestens einem Folgewert, die mit Hilfe des Fragebogens EORTC QLQ-C30 erfasst wurden.

Tabelle 4-29: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**Mittelwert (SD) (Punkte)**	**N**	**Mittelwert (SD) (Punkte)**
Fatigue	467	26,4 (22,12)	461	25,3 (21,19)
Übelkeit und Erbrechen	467	4,9 (12,70)	461	5,0 (12,41)
Schmerz	467	18,5 (23,38)	461	16,5 (21,92)
Dyspnoe	465	13,8 (21,37)	460	14,0 (21,56)
Schlaflosigkeit	467	19,3 (25,56)	461	20,7 (25,47)
Appetitminderung	467	14,4 (23,75)	461	13,5 (22,90)
Obstipation	466	11,8 (22,32)	460	12,6 (21,68)
Diarrhö	465	7,7 (16,59)	460	5,9 (13,81)
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; N: Anzahl der Patienten mit einem Baseline-Wert; SD: Standard Deviation (Standardabweichung) Quelle: SCO:SSiS 2018				

Die mittleren Baseline-Werte der Symptome, gemessen auf den Skalen des Fragebogens EORTC QLQ-C30 (Symptome), sind in Tabelle 4-29 dargestellt. Die Mittelwerte sowie die Standardabweichungen der beiden Therapiearme Lenvatinib und Sorafenib sind vergleichbar und lassen auf eine gute Randomisierung schließen.

Hauptanalyse

Im Folgenden werden die Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Symptome, gemessen auf den Skalen des Fragebogens EORTC QLQ-C30, dargestellt. Diese werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

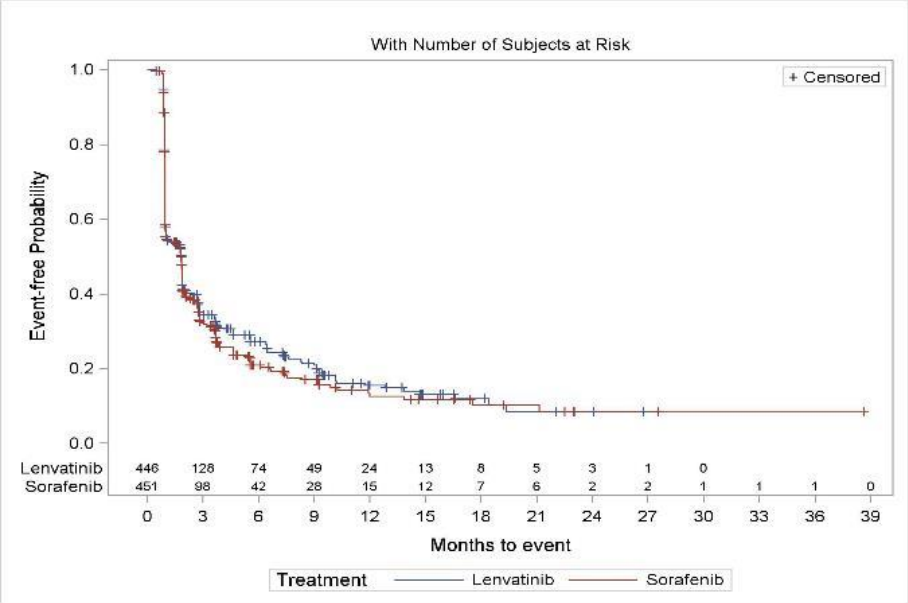


Abbildung 4-7: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

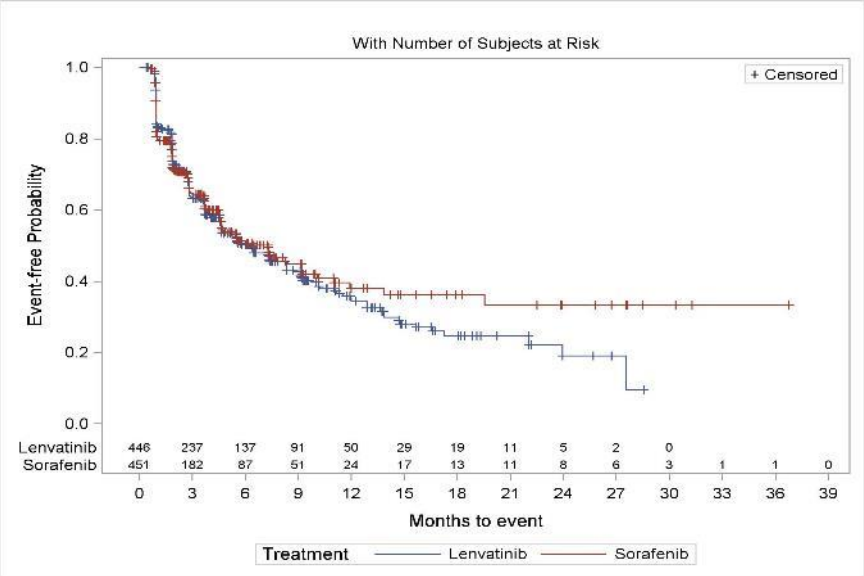


Abbildung 4-8: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Übelkeit** und **Erbrechen** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

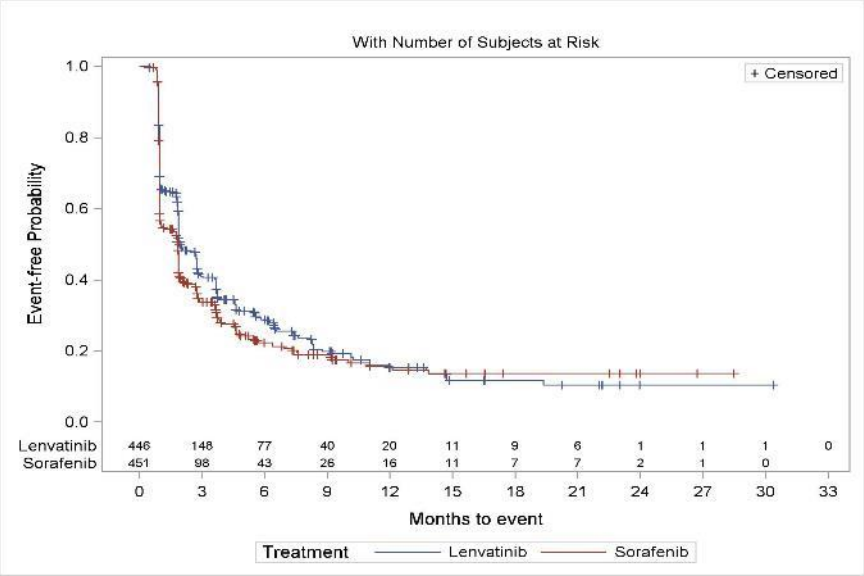


Abbildung 4-9: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

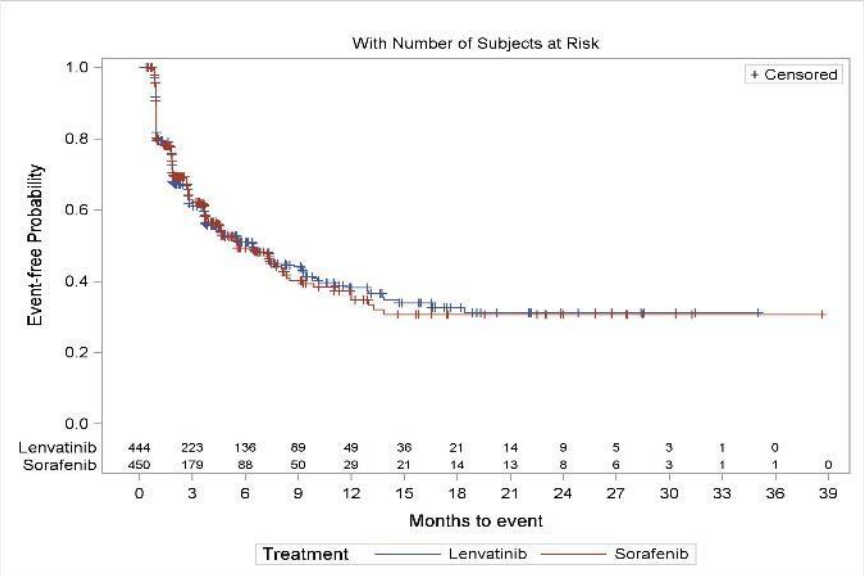


Abbildung 4-10: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

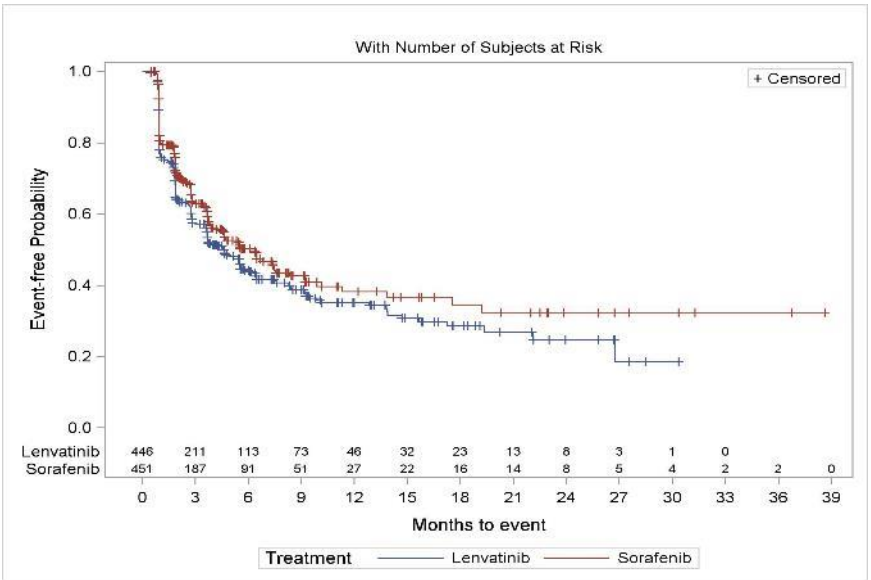


Abbildung 4-11: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

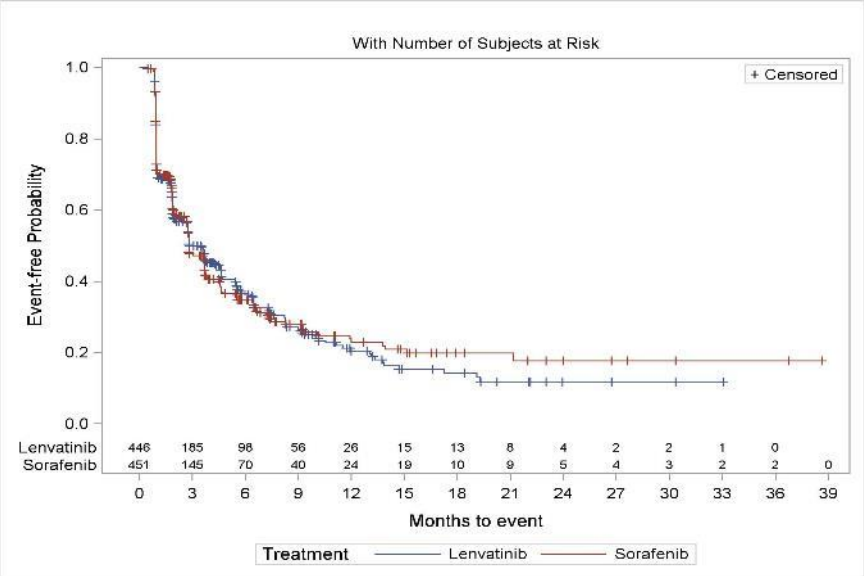


Abbildung 4-12: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

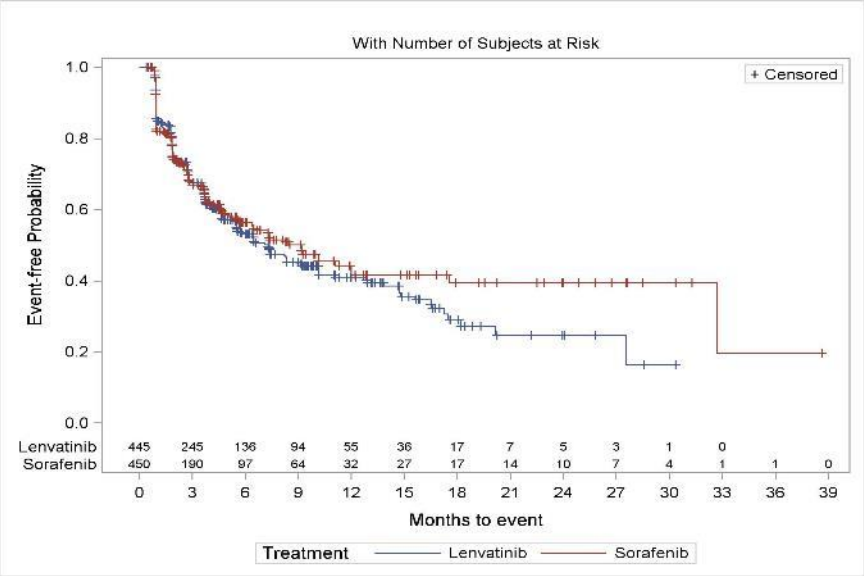


Abbildung 4-13: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Obstipation** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

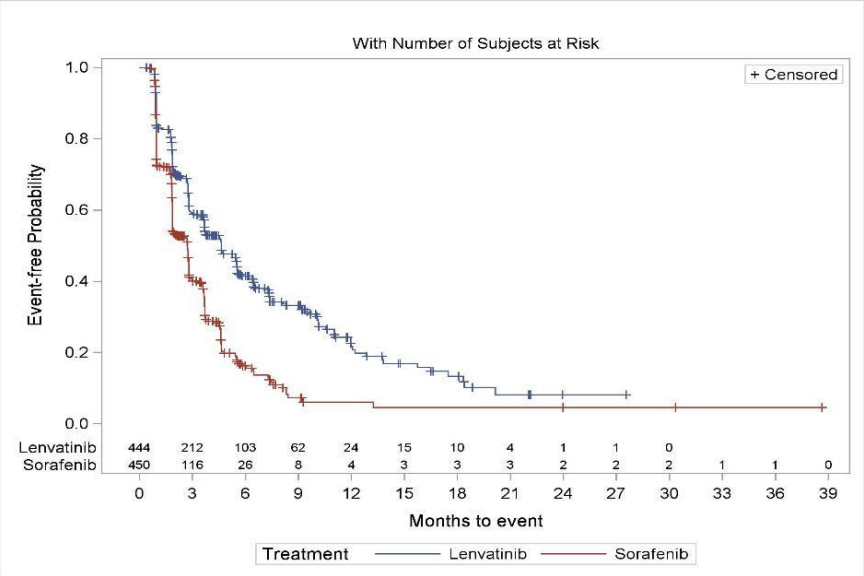


Abbildung 4-14: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

Tabelle 4-30: Hauptanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Fatigue	338/446 (75,8)	1,8 [1,05; 1,87]	330/451 (73,2)	1,8 [1,02; 1,87]	0,93 [0,797; 1,085] 0,3592
Übelkeit und Erbrechen	241/446 (54,0)	6,4 [4,60; 8,31]	189/451 (41,9)	7,3 [4,63; 9,20]	1,04 [0,857; 1,263] 0,6894
Schmerz	321/446 (72,0)	1,9 [1,87; 2,76]	319/451 (70,7)	1,8 [1,12; 1,87]	0,81 [0,690; 0,949] 0,0091
Dyspnoe	230/444 (51,8)	6,4 [4,57; 8,28]	202/450 (44,9)	5,6 [4,57; 8,02]	0,97 [0,797; 1,170] 0,7217
Schlaflosigkeit	251/446 (56,3)	4,6 [3,68; 5,55]	199/451 (44,1)	6,1 [4,53; 7,56]	1,16 [0,958; 1,398] 0,1293

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Appetitminderung	298/446 (66,8)	2,8 [2,76; 3,78]	268/451 (59,4)	2,8 [2,76; 3,68]	1,02 [0,858; 1,201] 0,8575
Obstipation	219/445 (49,2)	7,3 [5,49; 9,20]	175/450 (38,9)	9,2 [6,37; 11,99]	1,06 [0,868; 1,301] 0,5537
Diarrhö	266/444 (59,9)	4,6 [3,71; 5,52]	310/450 (68,9)	2,7 [1,87; 2,79]	0,52 [0,440; 0,623] <0,0001

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung" als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

HR: Hazard Ratio; IxRS: Interaktives Voice and Web Response Systems; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert; KI: Konfidenzintervall; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

Die Hauptanalyse der Ergebnisse des Fragebogens EORTC QLQ-C30 zeigte für die den Patienten beeinträchtigenden und belastenden Symptome Schmerz und Diarrhö eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell zeigte Lenvatinib gegenüber Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Schmerz eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,81; 95%-KI [0,690; 0,949]; p=0,0091). Diese statistisch signifikante Überlegenheit konnte auch für das den Patienten beeinträchtigende und belastende Symptom Diarrhö gezeigt werden (HR=0,52; 95%-KI [0,440; 0,623]; p<0,0001) (SCO:SSiS 2018).

Bezüglich des Symptoms Diarrhö betrug der absolute Vorteil in der medianen Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte 1,9 Monate zugunsten von Lenvatinib (4,6 Monate; 95%-KI [3,71; 5,52]) im Vergleich zu Sorafenib (2,7 Monate; 95%-KI [1,87; 2,79]) (SCO:SSiS 2018). Die mediane Zeit für das Symptom Schmerz betrug 1,9 Monate (95%-KI [1,87; 2,76]) für Lenvatinib im Vergleich zu 1,8 Monate (95%-KI [1,12; 1,87]) für Sorafenib. Die Kaplan-Meier-Kurve für das Symptom Schmerz zeigt eine Trennung der Kurven und einen Vorteil für Lenvatinib über den aufgrund genügend vorliegender Datenpunkte interpretierbaren Beobachtungszeitraum, wobei die Kurven beim Median eng beieinanderliegen (Abbildung 4-9).

Insgesamt belegen die Ergebnisse der Hauptanalyse eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Symptome Schmerz und Diarrhö. Hinsichtlich der Symptome Fatigue, Übelkeit und Erbrechen, Dyspnoe, Schlaflosigkeit, Appetitminderung und Obstipation zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Sensitivitätsanalyse

Im Folgenden werden die Ergebnisse für die mittlere Änderung zur Baseline im Score der Symptome, gemessen auf den Skalen des EORTC QLQ-C30 (in Punkten), dargestellt. Diese Ergebnisse werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

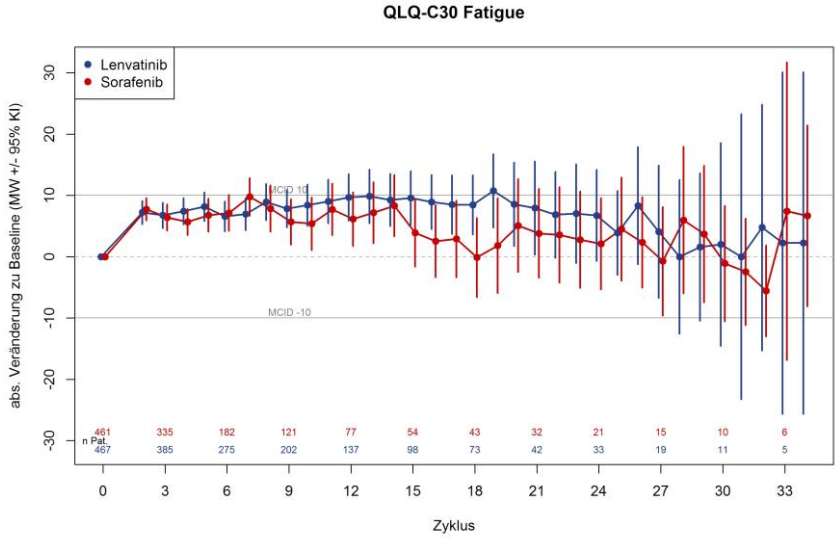


Abbildung 4-15: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

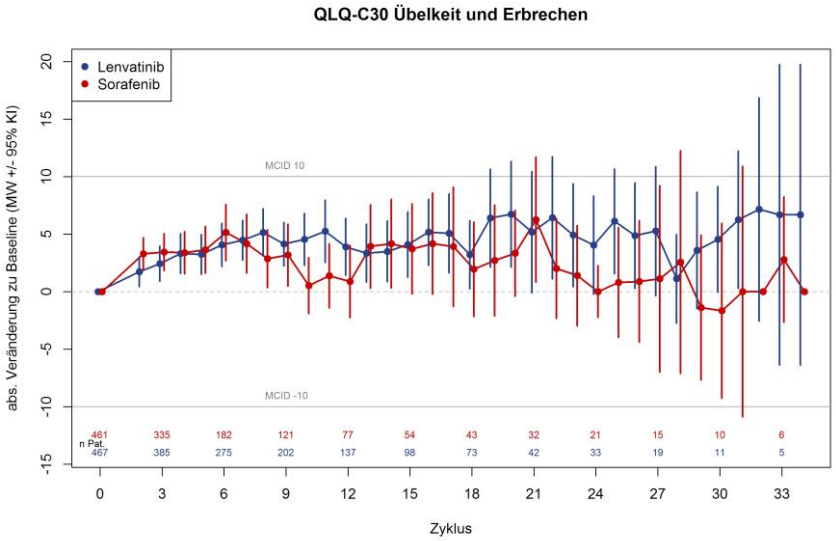


Abbildung 4-16: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Übelkeit und Erbrechen** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

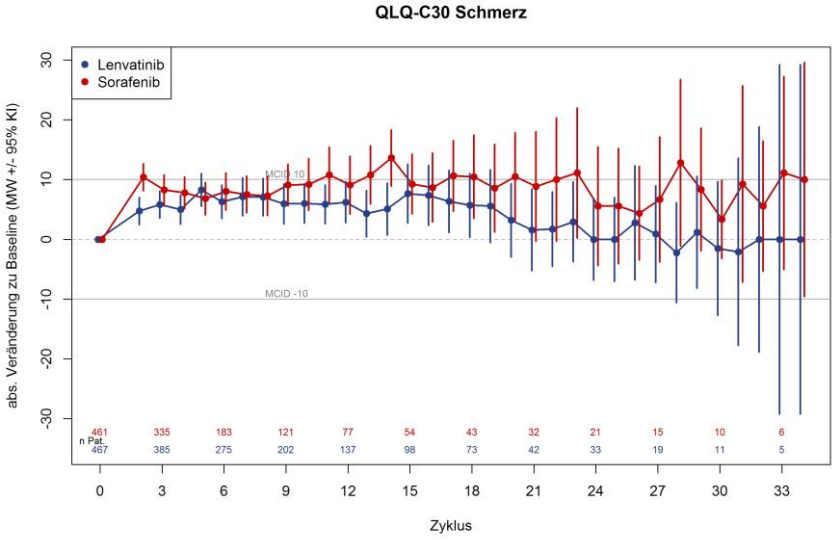


Abbildung 4-17: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

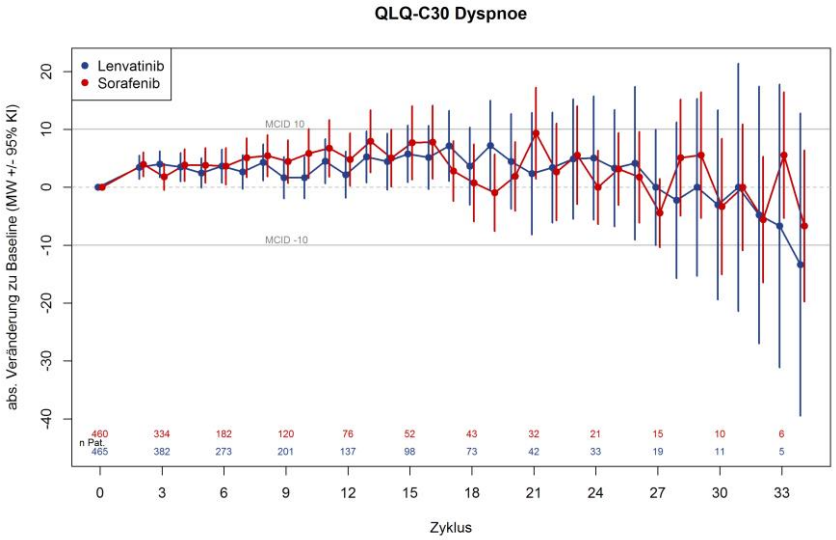


Abbildung 4-18: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

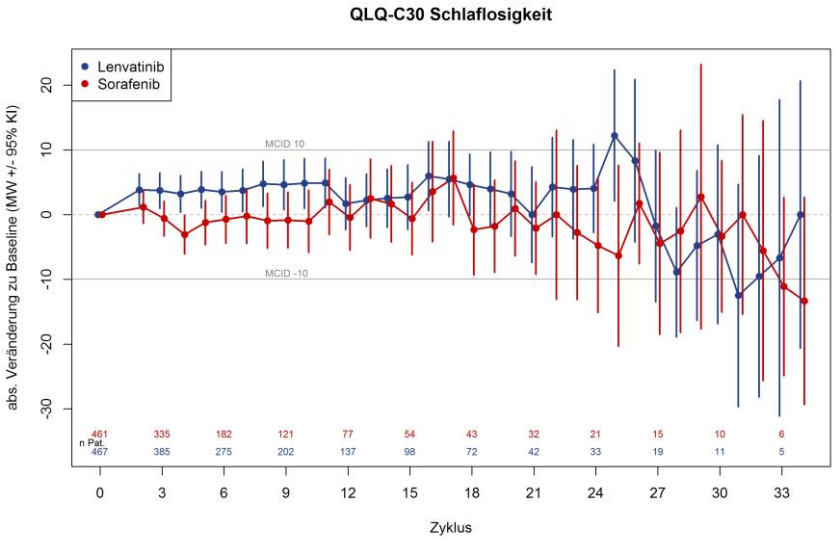


Abbildung 4-19: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

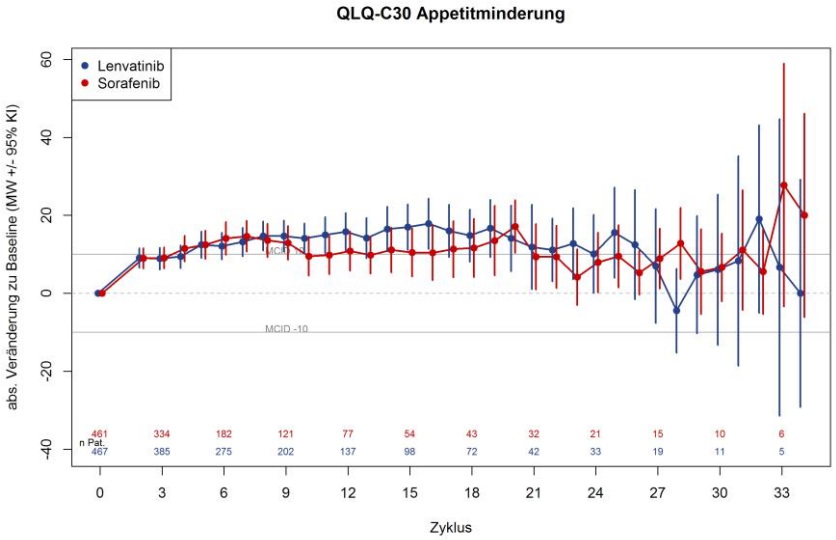


Abbildung 4-20: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

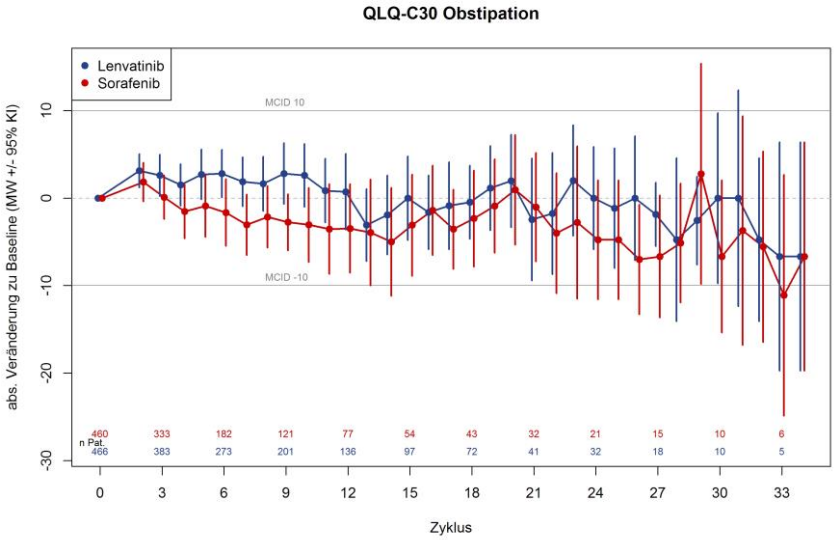


Abbildung 4-21: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Obstipation** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

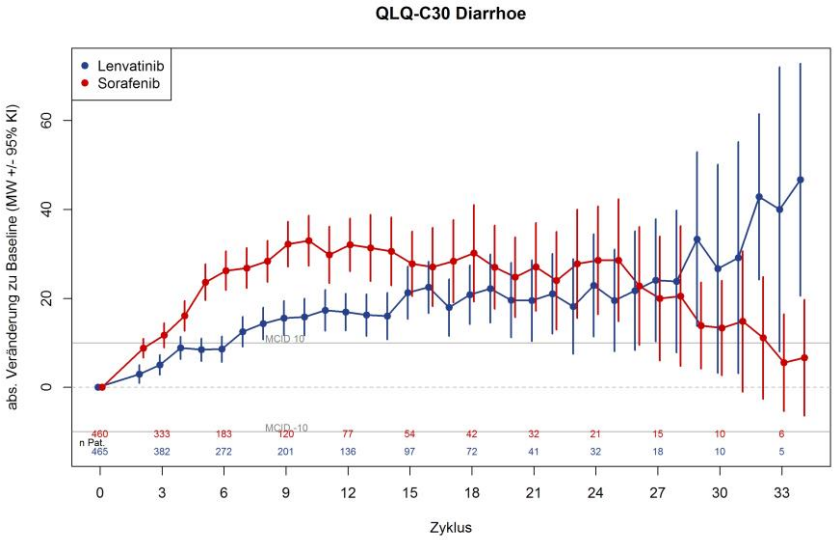


Abbildung 4-22: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2.3 EORTC QLQ-HCC18 (Symptome)– RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-31: Operationalisierung des Endpunkts EORTC QLQ-HCC18 (Symptome)

Studie	Operationalisierung
304	Der Fragebogen EORTC QLQ-HCC18 (Symptome) ist ein HCC-spezifischer Fragebogen mit Fragen bezüglich der Symptome Fatigue, Ikterus, Schmerz, Fieber und Abdominale Schwellung. Die einzelnen Items werden auf einer Skala von 1 bis 4 bewertet, welche wiederum linear in eine Skala von 0 bis 100 umgewandelt wird. Der Wert 100 stellt dabei den schlechtesten Status dar (Chie et al. 2012). Verändert sich bei einem Patienten der Wert auf einer der Skalen um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen. Die Analysen des Endpunkts EORTC QLQ-HCC18 (Symptome) wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. <u>Hauptanalyse:</u> Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurden mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt. <u>Sensitivitätsanalyse:</u> Für die Analyse der mittleren Veränderung zur Baseline im Score des jeweiligen Symptoms wurden der Mittelwert zu Baseline und der LS-Mittelwert der Änderung zur Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zur Baseline analysiert. Bei der Analyse der mittleren Veränderung zur Baseline im Score des jeweiligen Symptoms handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.
EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; LS: Least Square	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-32: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-HCC18 (Symptome) in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-HCC18 beschrieben.

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95% (Tabelle 4-33). Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-HCC18 als hoch eingestuft (Tabelle 4-32).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Rücklaufquoten und die Ergebnisse der Haupt- und Sensitivitätsanalyse des Fragebogens EORTC QLQ-HCC18 (Symptome) für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt.

Rücklaufquoten

Tabelle 4-33: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**n (%)**	**N**	**n (%)**
Fatigue	478	443 (92,7)	476	448 (94,1)
Ikterus	478	442 (92,5)	476	449 (94,3)
Schmerz	478	443 (92,7)	476	449 (94,3)
Fieber	478	443 (92,7)	476	449 (94,3)
Abdominale Schwellung	478	443 (92,7)	476	449 (94,3)
EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); ITT: Intention-to-treat; n: Anzahl der Patienten mit Baseline und mind. einem Folgewert; N: Anzahl der Patienten in der ITT-Population Quelle: SCO:SSiS 2018				

Die Rücklaufquoten des Fragebogens EORTC QLQ-HCC18 (Symptome) sind in Tabelle 4-33 dargestellt. Hierbei handelt es sich um den Anteil derjenigen Patienten in der ITT-Population mit einem Baseline-Wert und mindestens einem Folgewert, die mit Hilfe des Fragebogens EORTC QLQ-HCC18 erfasst wurden.

Tabelle 4-34: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**Mittelwert (SD) (Punkte)**	**N**	**Mittelwert (SD) (Punkte)**
Fatigue	463	23,3 (19,72)	458	21,8 (19,52)
Ikterus	462	8,4 (13,57)	459	7,8 (12,57)
Schmerz	463	13,9 (17,74)	459	14,8 (17,62)
Fieber	463	5,7 (12,12)	459	5,4 (12,58)
Abdominale Schwellung	463	13,5 (21,69)	459	12,6 (20,06)
EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); N: Anzahl der Patienten mit einem Baseline-Wert; SD: Standard Deviation (Standardabweichung) Quelle: SCO:SSiS 2018				

Die mittleren Baseline-Werte der Symptome gemessen auf den Skalen des Fragebogens EORTC QLQ-HCC18 (Symptome) sind in Tabelle 4-34 dargestellt. Die Mittelwerte sowie die Standardabweichungen der beiden Therapiearme Lenvatinib und Sorafenib sind vergleichbar und lassen auf eine gute Randomisierung schließen.

Hauptanalyse

Im Folgenden werden die Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Symptome, gemessen auf den Skalen des Fragebogens EORTC QLQ-HCC18, dargestellt. Diese werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

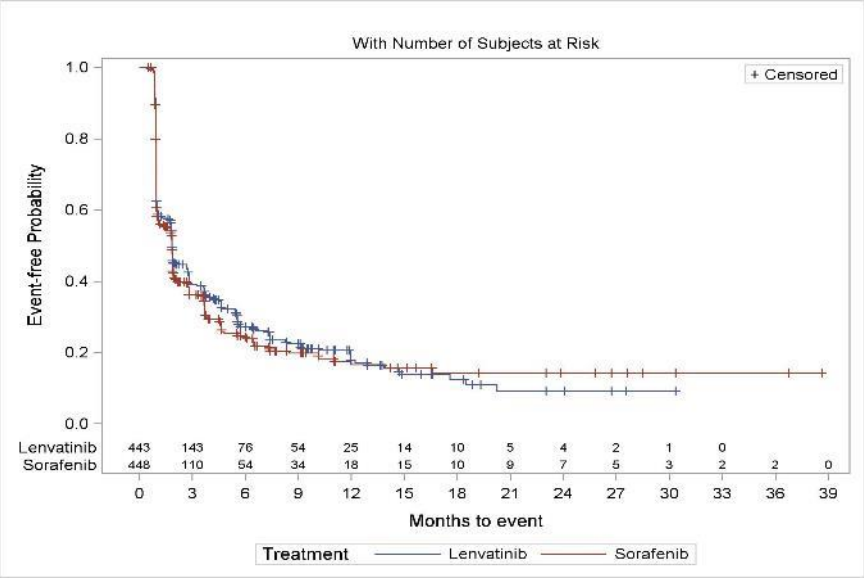


Abbildung 4-23: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

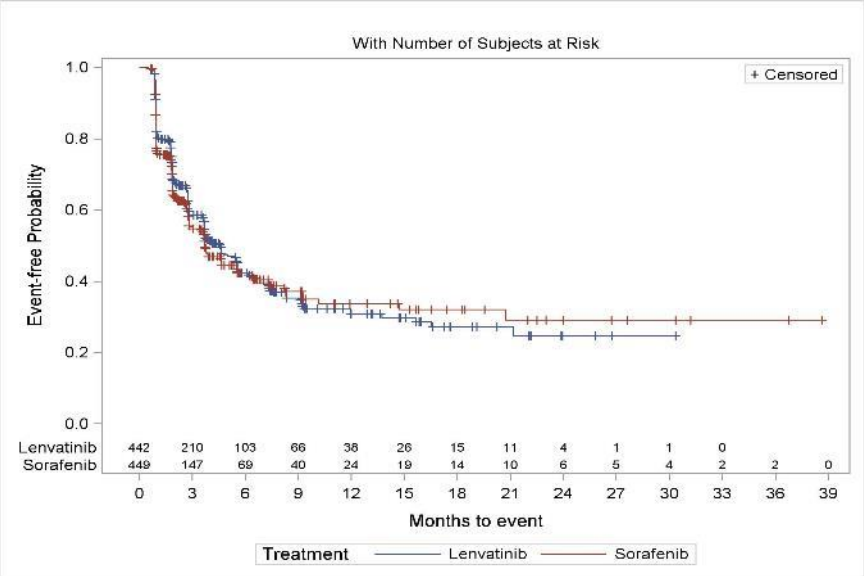


Abbildung 4-24: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ikterus** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

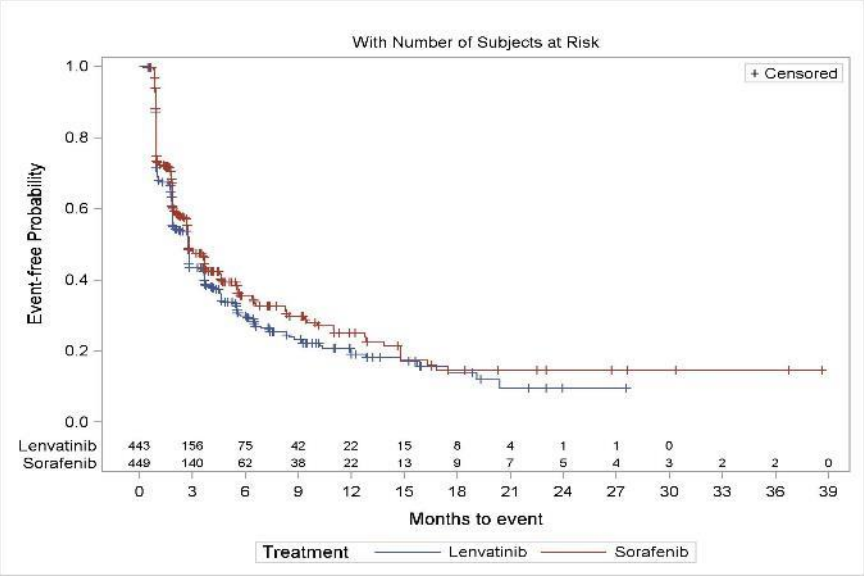


Abbildung 4-25: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

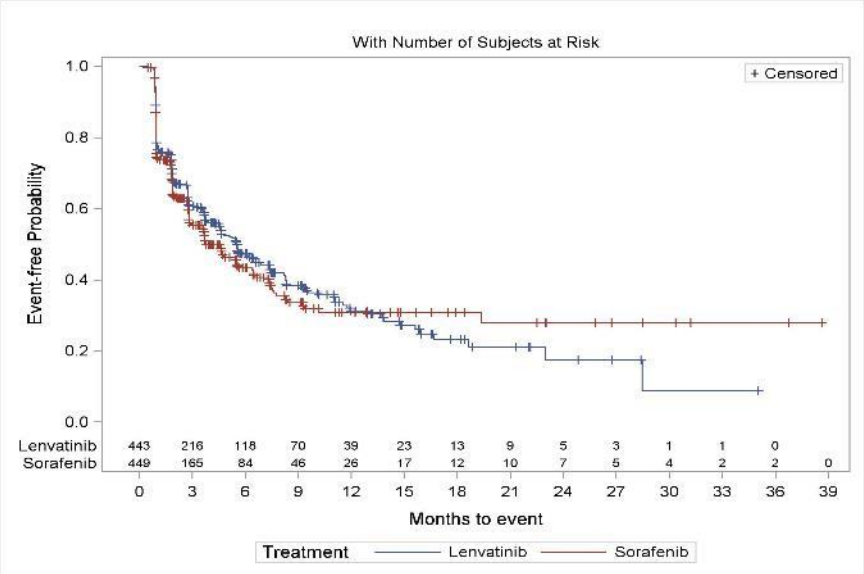


Abbildung 4-26: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

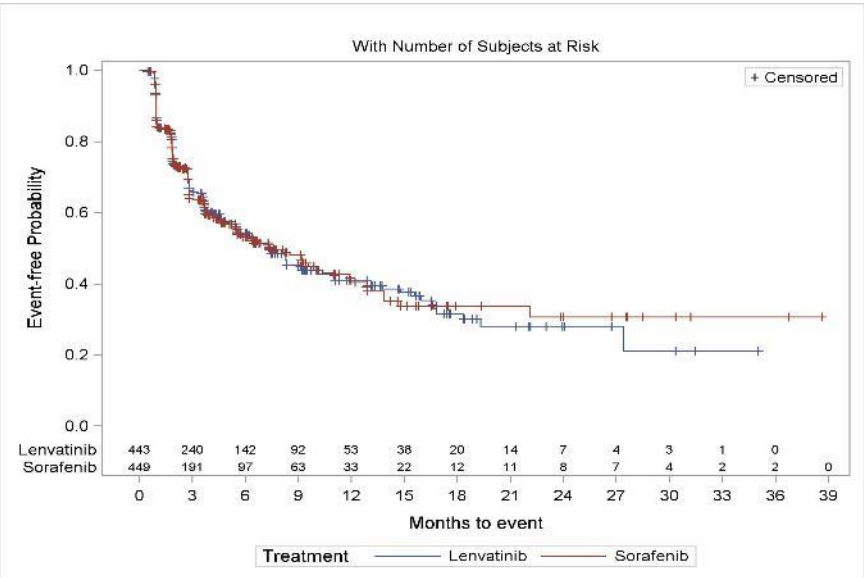


Abbildung 4-27: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Abdominale Schwellung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

Tabelle 4-35: Hauptanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Fatigue	328/443 (74,0)	1,8 [1,81; 1,91]	316/448 (70,5)	1,8 [1,77; 1,87]	0,94 [0,802; 1,097] 0,4216
Ikterus	246/442 (55,7)	4,6 [3,68; 5,55]	219/449 (48,8)	3,7 [2,89; 4,67]	0,95 [0,790; 1,144] 0,5934
Schmerz	303/443 (68,4)	2,8 [2,00; 2,79]	258/449 (57,5)	2,8 [2,73; 3,68]	1,14 [0,961; 1,349] 0,1315
Fieber	244/443 (55,1)	5,5 [4,57; 6,47]	231/449 (51,4)	3,7 [3,61; 5,52]	0,90 [0,752; 1,083] 0,2705
Abdominale Schwellung	218/443 (49,2)	7,4 [5,52; 9,20]	189/449 (42,1)	7,4 [5,49; 10,15]	0,98 [0,805; 1,197] 0,8549

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert

Quelle: SCO:SSiS 2018

Die Ergebnisse der Hauptanalyse des Fragebogens EORTC QLQ-HCC18 zeigten bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Symptome Fatigue, Ikterus, Fieber und Abdominale Schwellung, gemessen auf den jeweiligen Skalen des EORTC QLQ-HCC18, einen numerischen Vorteil und bezüglich des Symptoms Schmerz einen numerischen Nachteil von Lenvatinib im Vergleich zu Sorafenib (Tabelle 4-35). Die sich hier ergebene Diskrepanz bezüglich des Items Schmerz gemessen auf der Skala des EORTC QLQ-C30, kann mit den unterschiedlichen Fragen der beiden Instrumente EORTC QLQ-C30 und QLQ-HCC18 begründet werden. Während die Fragen des EORTC QLQ-C30 sich eher allgemein auf Schmerzen beziehen („Did pain interfere with your daily activities?“, „Have you

had pain?"), beziehen sich die Fragen beim EORTC QLQ-HCC18 auf lokalisationsspezifischen Schmerz („Have you had pain in your shoulder?"; "Have you had abdominal pain?").

Insgesamt zeigten die Ergebnisse der Hauptanalyse für Lenvatinib vergleichbare Therapieerfolge zu Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Symptoms, gemessen auf den Skalen des EORTC QLQ-HCC18 Fragebogens.

Sensitivitätsanalyse

Im Folgenden werden die Ergebnisse für die mittlere Änderung zur Baseline im Score der Symptome, gemessen auf den Skalen des EORTC QLQ-HCC18 (in Punkten), dargestellt. Diese werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

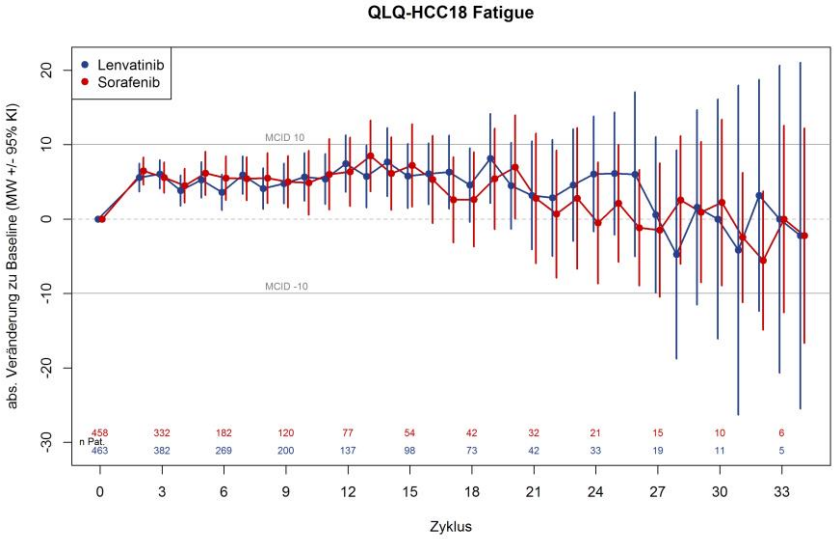


Abbildung 4-28: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

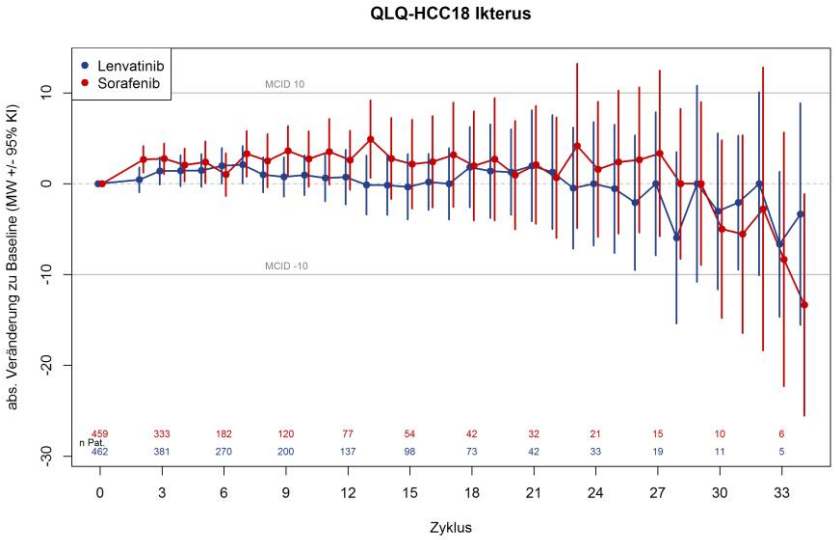


Abbildung 4-29: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Ikterus** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

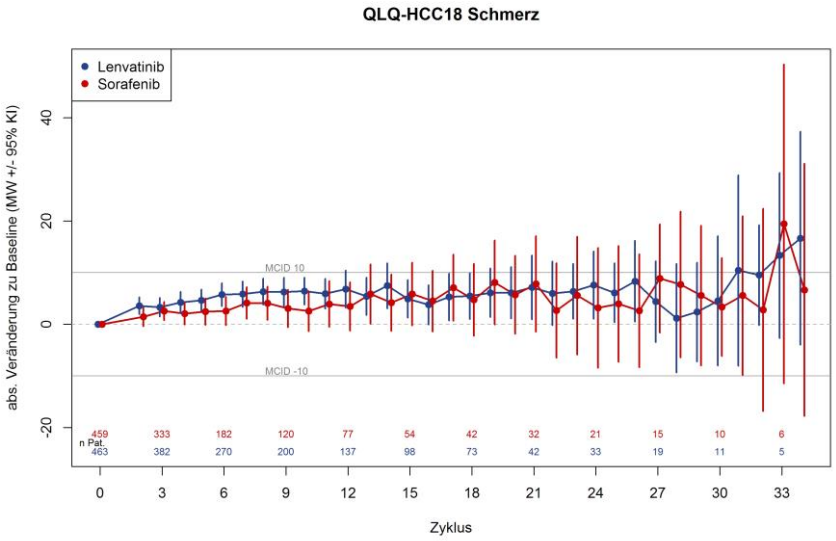


Abbildung 4-30: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

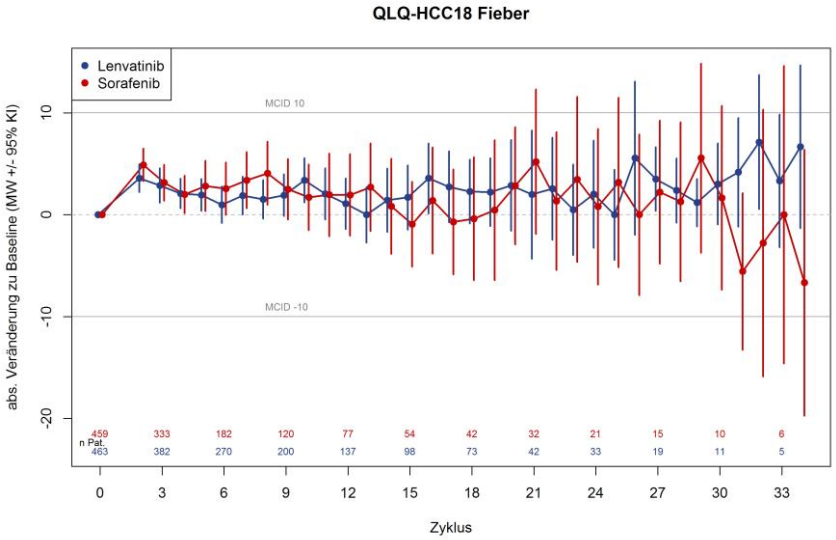


Abbildung 4-31: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

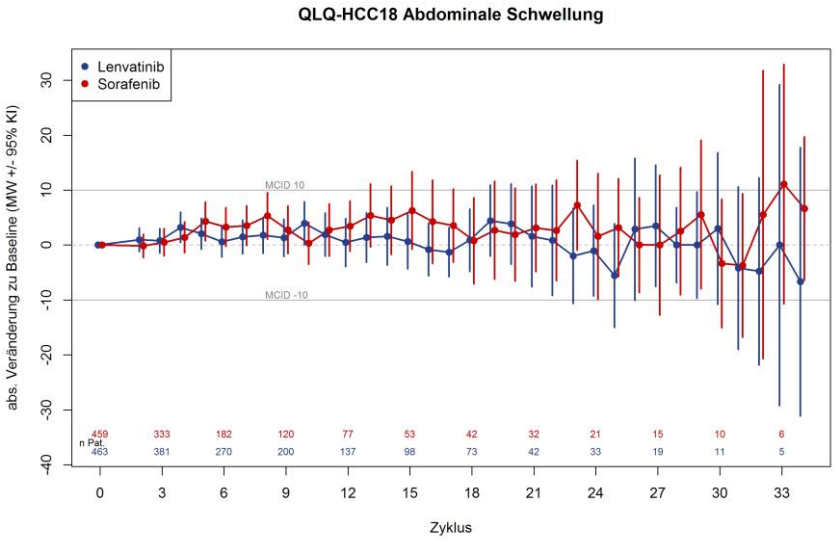


Abbildung 4-32: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Abdominale Schwellung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2.4 EQ-5D VAS – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-36: Operationalisierung des Endpunkts EQ-5D VAS

Studie	Operationalisierung
304	Der Fragebogen EQ-5D ist ein von der internationalen Forschungsgruppe EuroQoL entwickeltes, verlässliches und validiertes Messinstrument für den Gesundheitszustand eines Patienten. Er besteht aus einem deskriptiven Teil und einer visuellen Analogskala (engl. Visual Analogue Scale, VAS), die zusätzlich in den Fragebogen aufgenommen wurde, um den aktuellen Gesundheitszustand auf einer Skala von 0 bis 100 zu messen, wobei 0 für den "schlechtesten denkbaren Gesundheitszustand" und 100 für den "besten denkbaren Gesundheitszustand" steht (Pickard et al. 2012). Der Fragebogen wurde und wird derzeit in zahlreichen Studien mit Krebspatienten sowie speziell bei Patienten mit HCC eingesetzt (Pickard et al. 2012; Teckle et al. 2013). Eine Veränderung auf der VAS eines Patienten um mindestens 7 oder 10 mm im Vergleich zum Ausgangswert wurde hierbei als Response definiert (Pickard et al. 2007). Die Analysen des Endpunkts des EQ-5D VAS wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. <u>Hauptanalyse:</u> Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert.

	Die mediane Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurden mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt. Sensitivitätsanalyse: Für die Analyse der mittleren Veränderung zur Baseline auf der VAS wurden der Mittelwert zu Baseline und der LS-Mittelwert der Änderung zur Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zur Baseline analysiert. Bei der Analyse der mittleren Veränderung zur Baseline auf der VAS handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.
EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; HR: Hazard Ratio; IxRS: Interaktives Voice and Web Response Systems; ITT: Intention-to-treat; KI: Konfidenzintervall; LS: Least Square; mm: Millimeter; VAS: visuelle Analogskala	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-37: Bewertung des Verzerrungspotenzials für den Endpunkt EQ-5D VAS in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den Endpunkt EQ-5D VAS beschrieben.

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95% (Tabelle 4-38). Eine ergebnisabhängige Berichterstattung kann ausgeschlossen

werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EQ-5D VAS als hoch eingestuft (Tabelle 4-37).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Rücklaufquoten und die Ergebnisse der Haupt- und Sensitivitätsanalyse des Fragebogens EQ-5D VAS für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt.

Rücklaufquoten

Tabelle 4-38: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des Fragebogens EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**n (%)**	**N**	**n (%)**
EQ-5D VAS	478	441 (92,3)	476	445 (93,5)
EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; ITT: Intention-to-treat; n: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert; N: Anzahl der Patienten in der ITT-Population Quelle: SCO:SSiS 2018				

Die Rücklaufquoten des Fragebogens EQ-5D sind in Tabelle 4-38 dargestellt. Hierbei handelt es sich um den Anteil derjenigen Patienten in der ITT-Population mit einem Baseline-Wert und mindestens einem Folgewert, die mit Hilfe des Fragebogens EQ-5D erfasst wurden.

Tabelle 4-39: Ergebnisse der Baseline-Werte auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**Mittelwert (SD) (mm)**	**N**	**Mittelwert (SD) (mm)**
EQ-5D VAS	462	73,5 (18,43)	456	74,6 (17,25)
EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; N: Anzahl der Patienten mit einem Baseline-Wert; SD: Standard Deviation (Standardabweichung) Quelle: SCO:SSiS 2018				

Die mittleren Baseline-Werte auf der EQ-5D VAS sind in Tabelle 4-39 aufgelistet. Die Mittelwerte sowie die Standardabweichungen der beiden Therapiearme Lenvatinib und Sorafenib sind vergleichbar und lassen auf eine gute Randomisierung schließen.

Hauptanalyse

Im Folgenden werden die Ergebnisse für Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS dargestellt. Diese werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

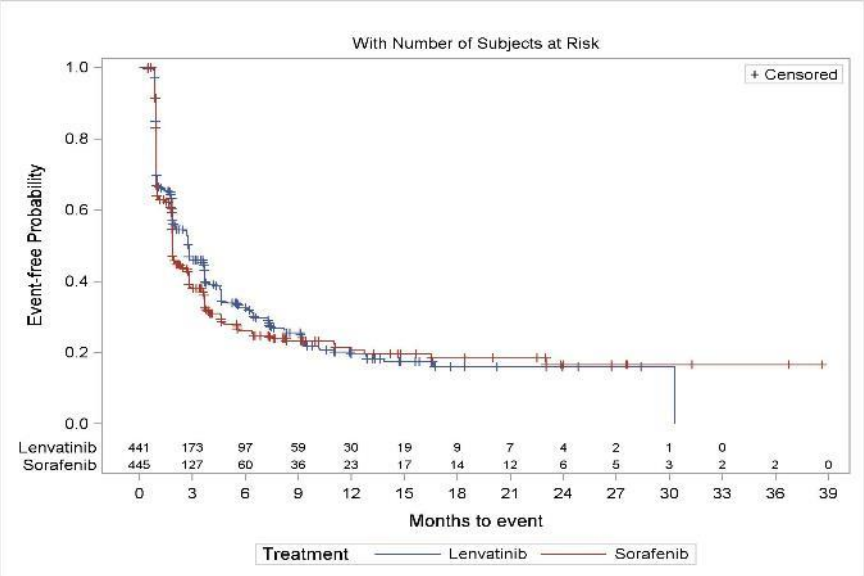


Abbildung 4-33: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 7 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel

EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale
Quelle: SCO:SSiS 2018

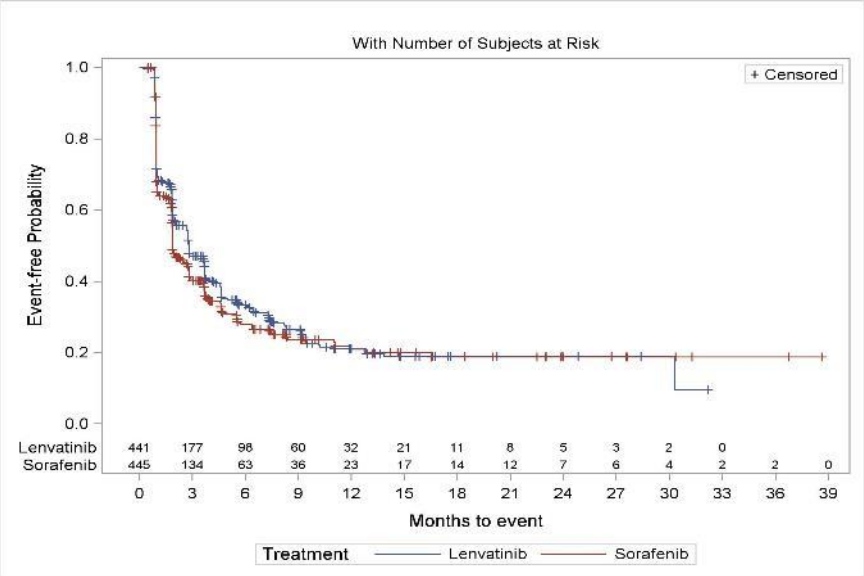


Abbildung 4-34: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel

EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale
Quelle: SCO:SSiS 2018

Tabelle 4-40: Hauptanalyse: Ergebnisse für Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
EQ-5D VAS, MID=7 mm	310/441 (70,3)	2,8 [2,10; 3,61]	303/445 (68,1)	1,9 [1,84; 2,04]	0,90 [0,765; 1,056] 0,1951
EQ-5D VAS, MID=10 mm	303/441 (68,7)	2,8 [2,69; 3,68]	295/445 (66,3)	1,9 [1,87; 2,56]	0,89 [0,758; 1,050] 0,1713

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; MID: Minimal Important Difference; mm: Millimeter; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert

	Lenvatinib		Sorafenib		Lenvatinib vs. Sorafenib
Endpunkt	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Quelle: SCO:SSiS 2018					

Die Hauptanalyse der Ergebnisse des Fragebogens EQ-5D zeigte bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS einen numerischen Vorteil von Lenvatinib im Vergleich zu Sorafenib (Tabelle 4-40).

Insgesamt zeigten die Ergebnisse der Hauptanalyse für Lenvatinib vergleichbare Therapieerfolge zu Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS.

Sensitivitätsanalyse

Im Folgenden werden die Ergebnisse für die mittlere Änderung zur Baseline für den Endpunkt EQ-5D VAS dargestellt. Diese werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

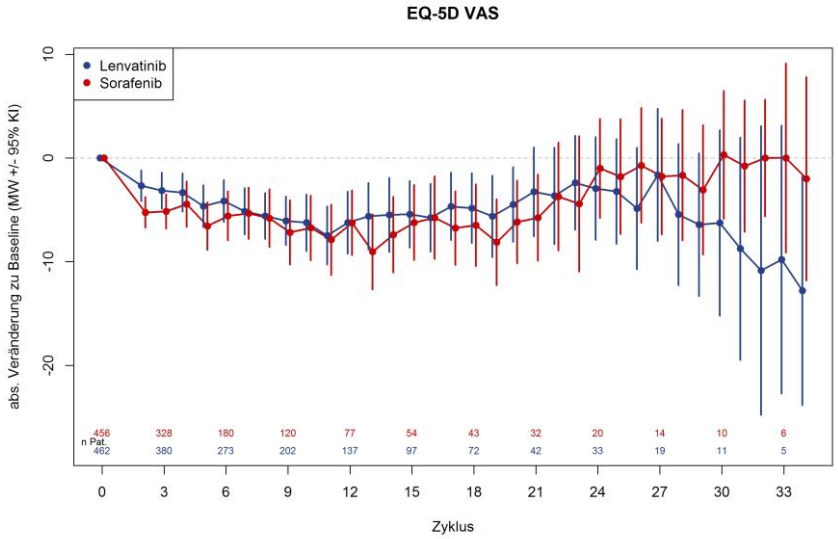


Abbildung 4-35: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel
EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale
Quelle: SCO:SSiS 2018

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse

durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.3 Endpunkt Gesundheitsbezogene Lebensqualität – RCT

4.3.1.3.1.3.1 EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-41: Operationalisierung des Endpunkts EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität)

Studie	Operationalisierung
304	Der Fragebogen EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) ist in der Onkologie ein häufig eingesetztes Instrument zur Bewertung der gesundheitsbezogenen Lebensqualität. Die gestellten Fragen beziehen sich dabei auf die Funktionen Physische Funktion, Rollenfunktionsniveau, Emotionale Funktion, Kognitive Funktion, Soziale Funktion, Finanzielle Schwierigkeiten und Globaler Gesundheitsstatus. Die einzelnen Items werden dabei auf einer Skala von 1 bis 4 bewertet bzw. 1 bis 7, für die Fragen zum globalen Gesundheitsstatus, die zur Auswertung in eine Skala von 0 bis 100 transformiert wird. Höhere Werte auf den Funktionsskalen und der Globalskala bedeuten eine höhere Lebensqualität (Giesinger et al. 2016). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen. Die Analysen des Endpunkts EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. <u>Hauptanalyse:</u> Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der jeweiligen Funktion wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der jeweiligen Funktion wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurden mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der jeweiligen Funktion wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt.

Sensitivitätsanalyse: Für die Analyse der mittleren Veränderung zur Baseline im Score der jeweiligen Funktion wurden der Mittelwert zu Baseline und der LS-Mittelwert der Änderung zur Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zur Baseline analysiert. Bei der Analyse der mittleren Veränderung zur Baseline im Score der jeweiligen Funktion handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; LS: Least Square

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-42: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den Endpunkt QLQ-C30 beschrieben.

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95% (Tabelle 4-43). Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-C30 als hoch eingestuft (Tabelle 4-42).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Rücklaufquoten und die Ergebnisse der Haupt- und Sensitivitätsanalyse des Fragebogens EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt.

Rücklaufquoten

Tabelle 4-43: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des Fragebogens EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**n (%)**	**N**	**n (%)**
Physische Funktion	478	446 (93,3)	476	451 (94,7)
Rollenfunktionsniveau	478	445 (93,1)	476	451 (94,7)
Emotionale Funktion	478	445 (93,1)	476	450 (94,5)
Kognitive Funktion	478	445 (93,1)	476	449 (94,3)
Soziale Funktion	478	445 (93,1)	476	450 (94,5)
Finanzielle Schwierigkeiten	478	443 (92,7)	476	448 (94,1)
Globaler Gesundheitsstatus	478	445 (93,1)	476	450 (94,5)

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ITT: Intention-to-treat; n: Anzahl der Patienten in der ITT-Population mit Baseline- und mind. einem Folgewert; N: Anzahl der Patienten in der ITT-Population

Quelle: SCO:SSiS 2018

Die Rücklaufquoten des Fragebogens EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) sind in Tabelle 4-43 dargestellt. Hierbei handelt es sich um den Anteil derjenigen Patienten in der ITT-Population mit einem Baseline-Wert und mindestens einem Folgewert, die mit Hilfe des Fragebogens EORTC QLQ-C30 erfasst wurden.

Tabelle 4-44: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**Mittelwert (SD) (Punkte)**	**N**	**Mittelwert (SD) (Punkte)**
Physische Funktion	467	84,8 (16,29)	461	85,6 (15,62)
Rollenfunktionsniveau	466	85,4 (21,23)	461	85,8 (22,03)
Emotionale Funktion	466	83,7 (16,94)	460	83,4 (16,80)

	Lenvatinib		Sorafenib	
Endpunkt	**N**	**Mittelwert (SD) (Punkte)**	**N**	**Mittelwert (SD) (Punkte)**
Kognitive Funktion	466	89,5 (14,26)	459	87,4 (17,14)
Soziale Funktion	466	82,7 (23,06)	460	82,8 (22,19)
Finanzielle Schwierigkeiten	464	27,8 (31,60)	459	26,7 (31,37)
Globaler Gesundheitsstatus	466	67,1 (20,47)	460	67,6 (20,52)

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; N: Anzahl der Patienten mit einem Baseline-Wert; SD: Standard Deviation (Standardabweichung)

Quelle: SCO:SSiS 2018

Die mittleren Baseline-Werte der Funktionen gemessen auf den Skalen des Fragebogens EORTC QLQ-C30 sind in Tabelle 4-44 aufgelistet. Die Mittelwerte sowie die Standardabweichungen der beiden Therapiearme Lenvatinib und Sorafenib sind vergleichbar und lassen auf eine gute Randomisierung schließen.

Hauptanalyse

Im Folgenden werden die Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Funktionen, gemessen auf den Skalen des EORTC QLQ-C30, dargestellt. Diese werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

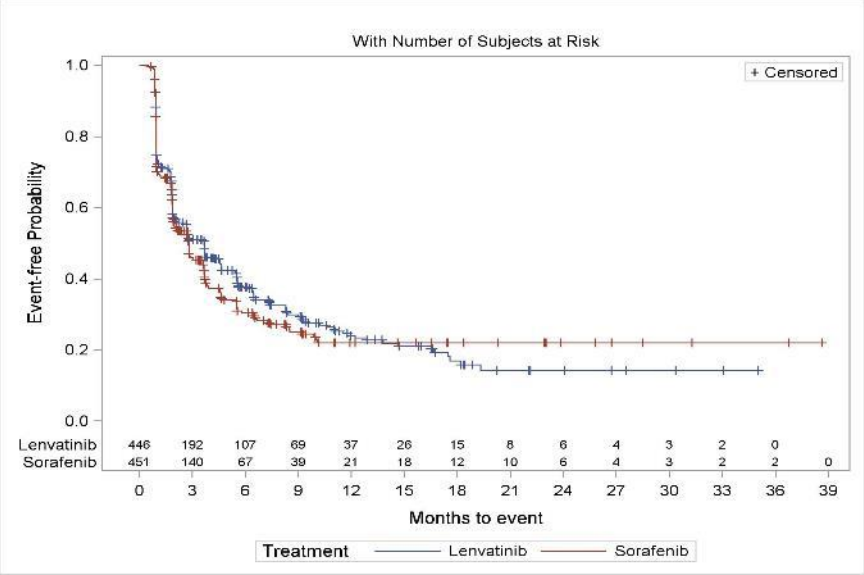


Abbildung 4-36: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

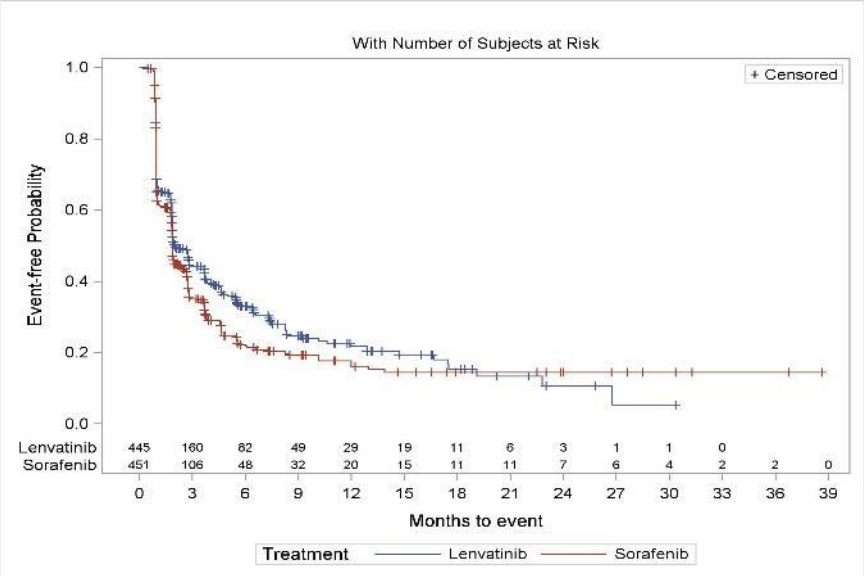


Abbildung 4-37: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Rollenfunktionsniveau** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

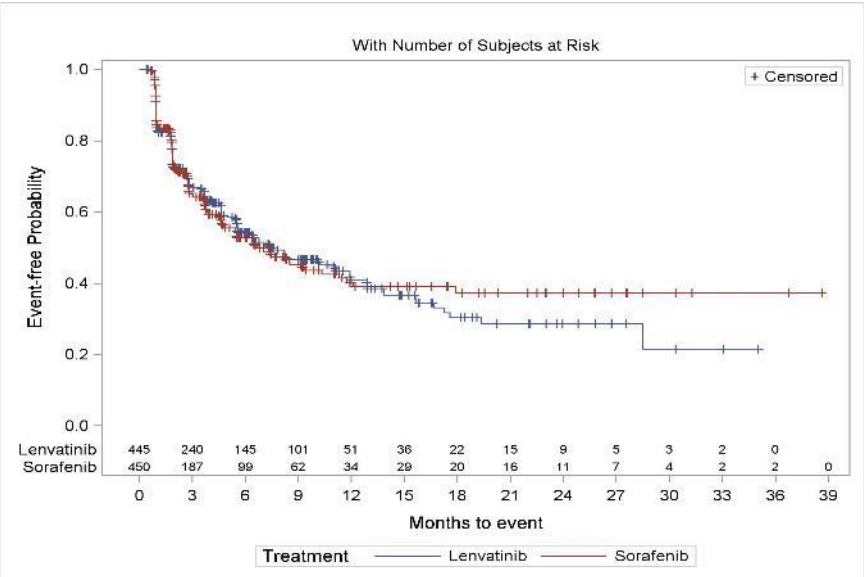


Abbildung 4-38: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

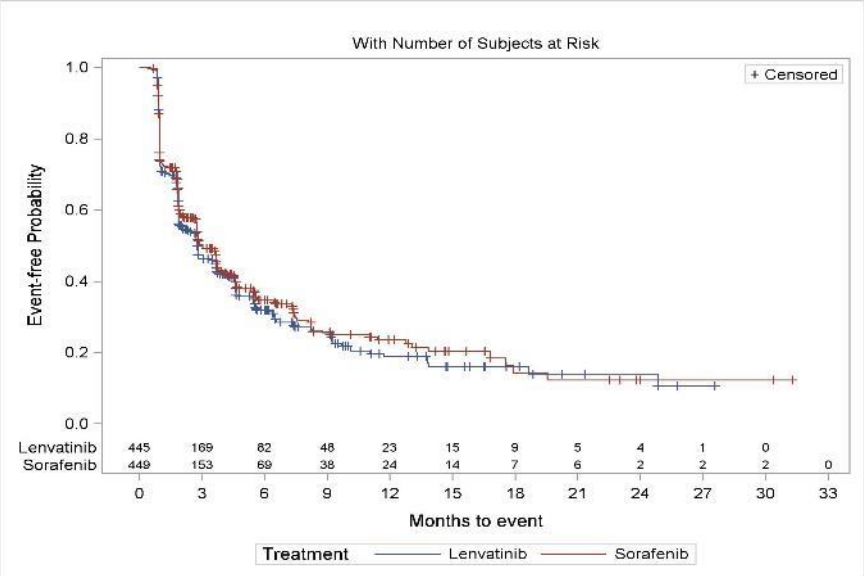


Abbildung 4-39: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

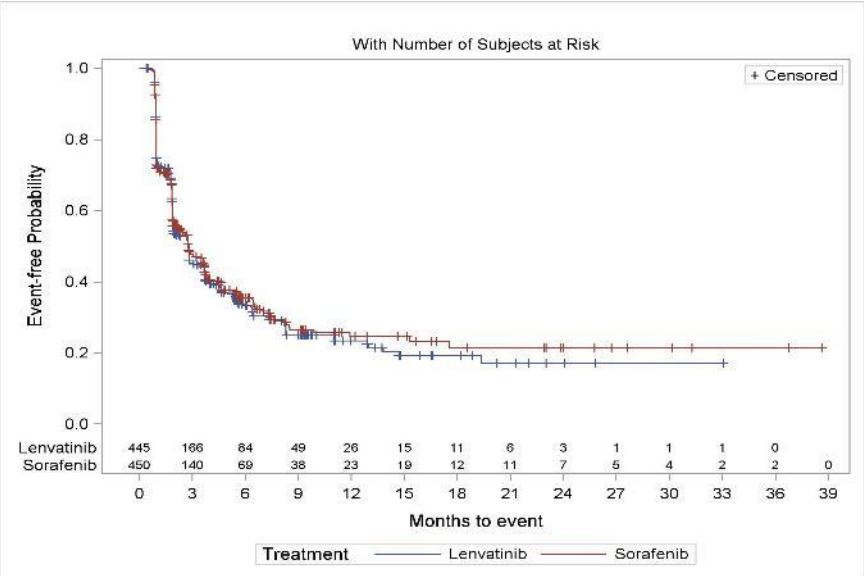


Abbildung 4-40: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

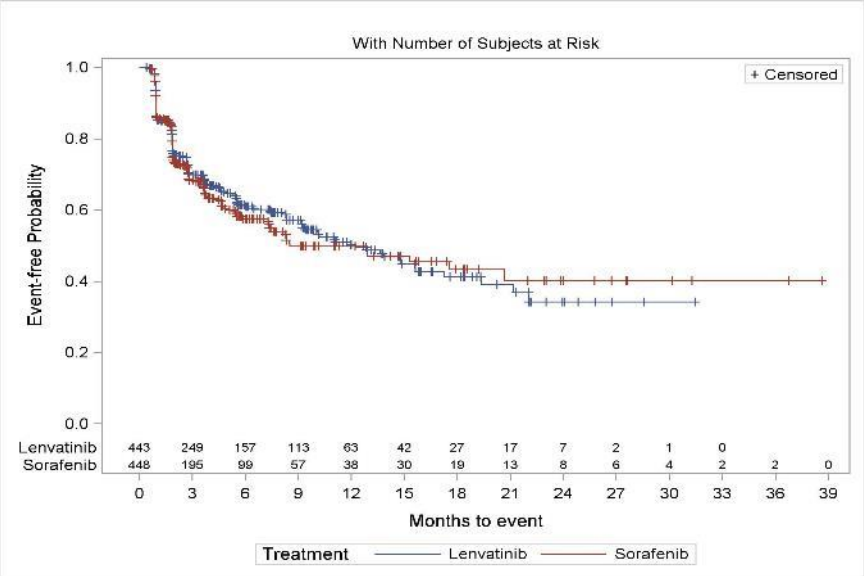


Abbildung 4-41: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Finanzielle Schwierigkeiten** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

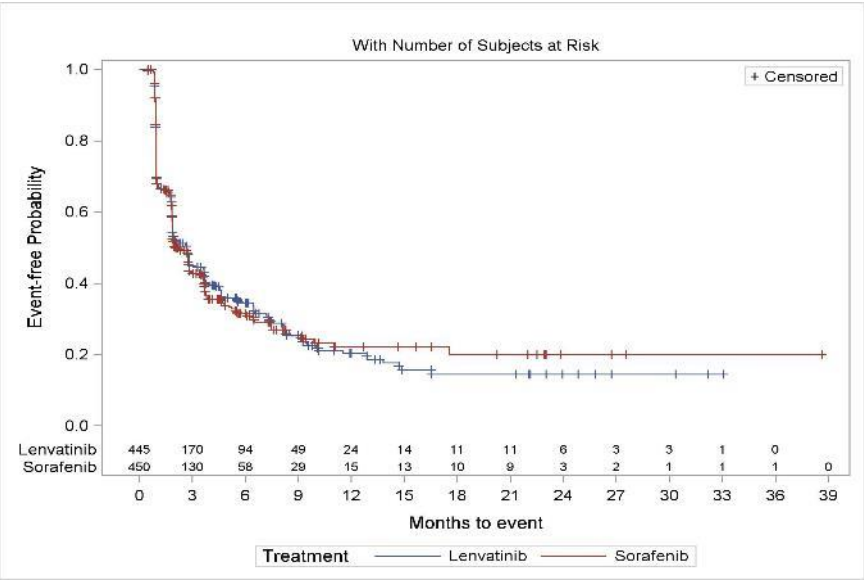


Abbildung 4-42: Hauptanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

Tabelle 4-45: Hauptanalyse: Ergebnisse für Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Physische Funktion	294/446 (65,9)	3,7 [2,69; 4,53]	277/451 (61,4)	2,8 [1,97; 3,61]	0,89 [0,756; 1,056] 0,1870
Rollenfunktionsniveau	302/445 (67,9)	2,0 [1,87; 2,83]	314/451 (69,6)	1,9 [1,84; 1,97]	0,82 [0,696; 0,962] 0,0148
Emotionale Funktion	216/445 (48,5)	7,4 [5,55; 11,07]	187/450 (41,6)	6,6 [4,80; 9,33]	1,01 [0,827; 1,235] 0,9142
Kognitive Funktion	301/445 (67,6)	2,8 [2,10; 3,65]	268/449 (59,7)	3,0 [2,76; 3,71]	1,09 [0,920; 1,291] 0,3167
Soziale Funktion	293/445 (65,8)	2,8 [1,91; 3,02]	264/450 (58,7)	2,8 [2,30; 3,68]	1,06 [0,899; 1,260] 0,4658
Finanzielle Schwierigkeiten	182/443 (41,1)	12,3 [9,23; 15,61]	166/448 (37,1)	8,5 [7,36; 20,67]	0,91 [0,735; 1,130] 0,3980
Globaler Gesundheitsstatus	307/445 (69,0)	2,7 [1,87; 2,79]	278/450 (61,8)	2,1 [1,87; 2,79]	0,99 [0,841; 1,171] 0,9247

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline-und mind. einem Folgewert

Quelle: SCO:SSiS 2018

Die Ergebnisse der Hauptanalyse des Fragebogens EORTC QLQ-C30 zeigte für das Rollenfunktionsniveau eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell zeigte sich bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des

Rollenfunktionsniveaus eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,82; 95%-KI [0,696; 0,962]; p=0,0148) (Tabelle 4-45).

Die mediane Zeit bis zur ersten Verschlechterung um 10 Punkte betrug dabei 2,0 Monate (95%-KI [1,87; 2,83]) für Lenvatinib im Vergleich zu 1,9 Monate (95%-KI [1,84; 1,97]) für Sorafenib (Tabelle 4-45). Die Kaplan-Meier-Kurve für das Rollenfunktionsniveau zeigt eine Trennung der Kurven und einen Vorteil für Lenvatinib über den aufgrund genügend vorliegender Datenpunkte interpretierbaren Beobachtungszeitraum, wobei die Kurven beim Median eng beieinanderliegen (Abbildung 4-37).

Bezüglich der Skalen Physische Funktion und Finanzielle Schwierigkeiten und dem globalen Gesundheitsstatus konnten numerische Vorteile und bezüglich der Funktionen Emotionale Funktion, Kognitive Funktion und Soziale Funktion numerische Nachteile in der Zeit bis zur ersten Verschlechterung um 10 Punkte im Score der jeweiligen Funktion beobachtet werden (Tabelle 4-45).

Insgesamt belegen die Ergebnisse der Hauptanalyse eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Rollenfunktionsniveaus. Hinsichtlich der Physischen Funktion, der Emotionalen Funktion, der Kognitiven Funktion, der Sozialen Funktion, den Finanziellen Schwierigkeiten und dem globalen Gesundheitsstatus zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Sensitivitätsanalyse

Im Folgenden werden die Ergebnisse für die mittlere Änderung zur Baseline im Score der Funktionen, gemessen auf den Skalen des Fragebogens EORTC QLQ-C30, dargestellt. Diese werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

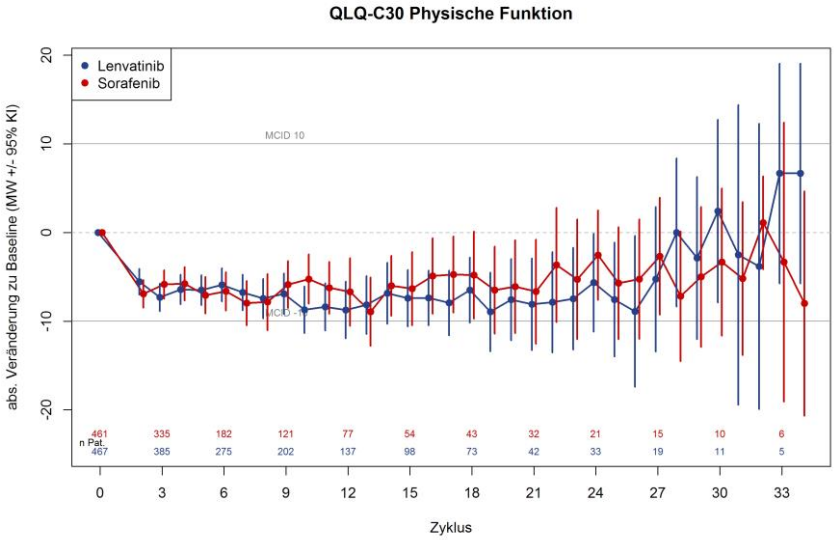


Abbildung 4-43: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

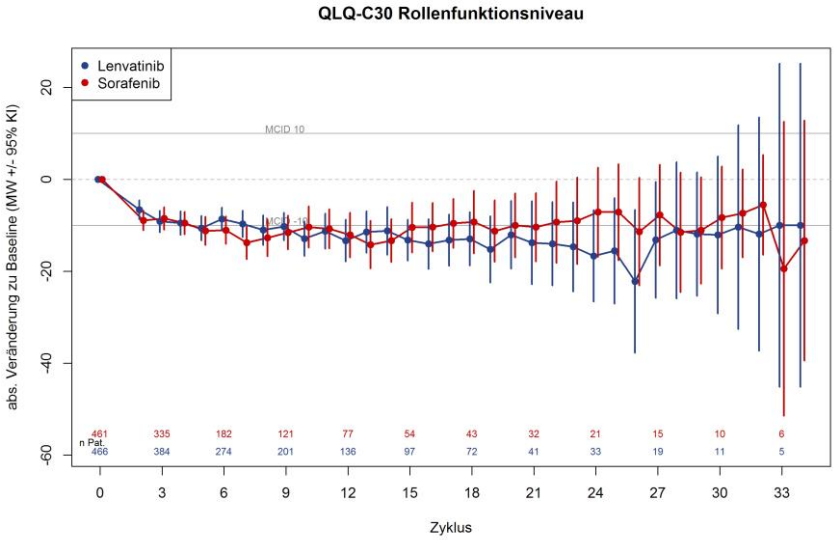


Abbildung 4-44: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Rollenfunktionsniveau** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

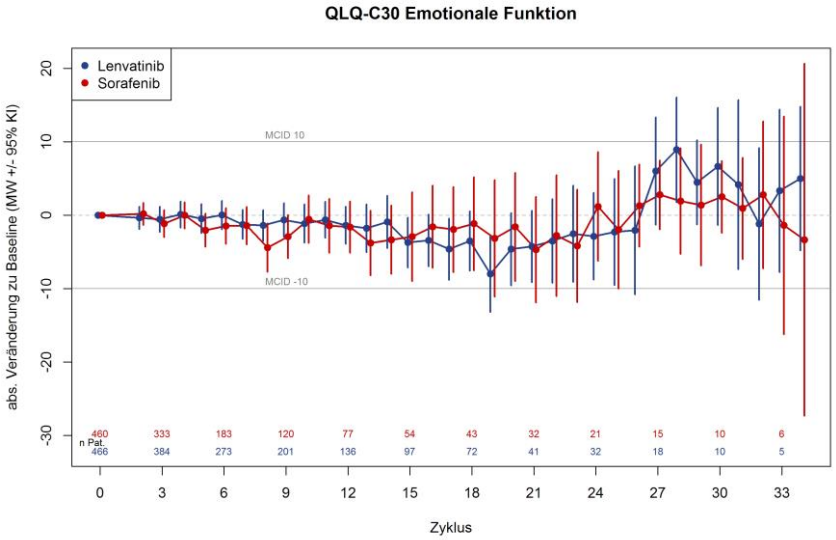


Abbildung 4-45: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

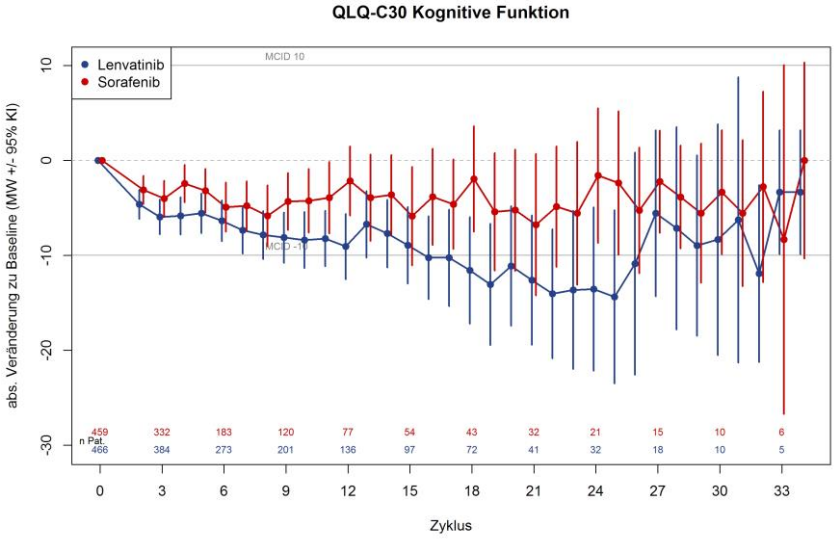


Abbildung 4-46: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

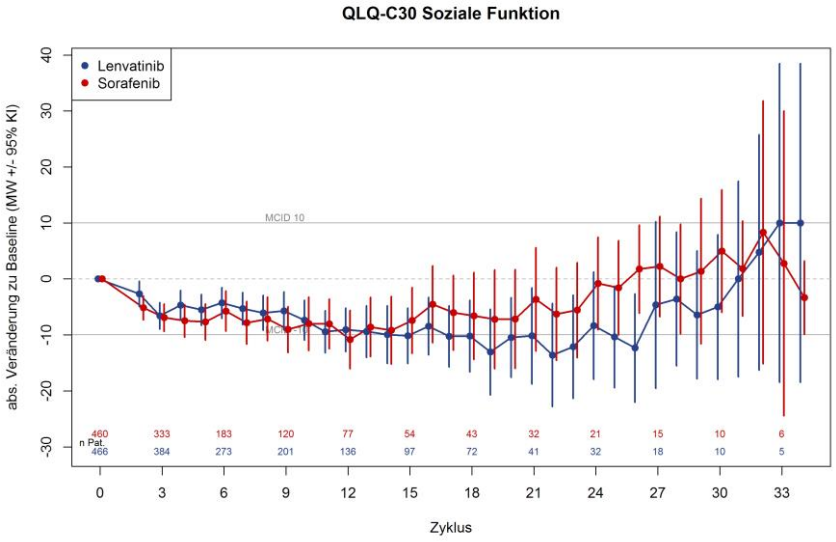


Abbildung 4-47: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

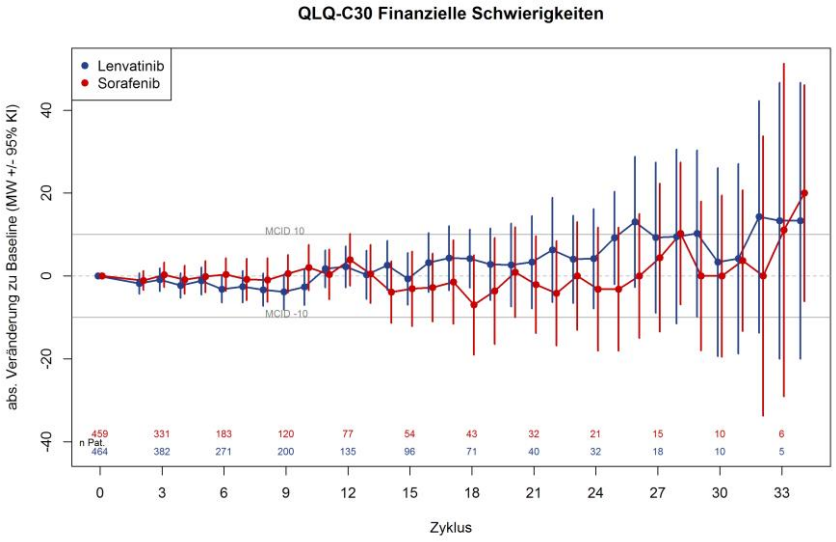


Abbildung 4-48: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Finanzielle Schwierigkeiten** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

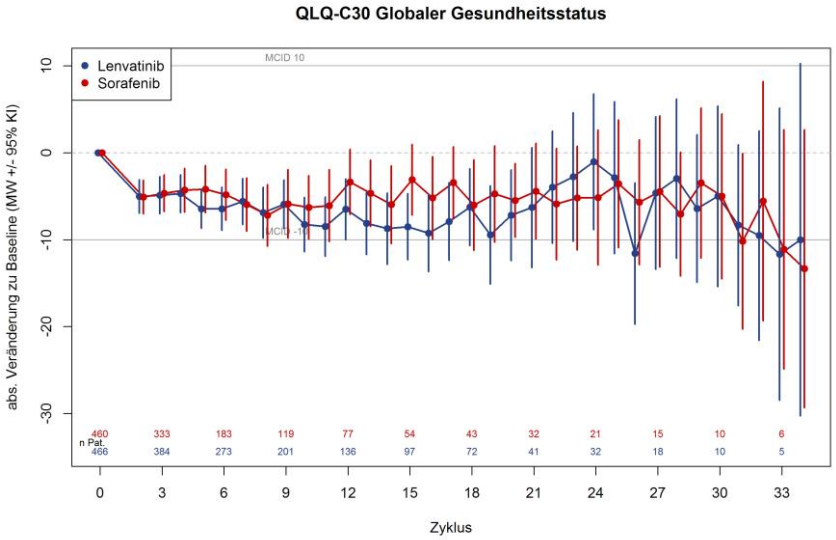


Abbildung 4-49 Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 aus Studie 304 mit dem zu bewertenden Arzneimittel
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.3.2 EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) – RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-46: Operationalisierung des Endpunkts EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität)

Studie	Operationalisierung
304	Der Fragebogen EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) ist ein HCC-spezifischer Fragebogen mit Fragen bezüglich der Items Körperbild, Ernährung und Sexualleben. Die einzelnen Items werden auf einer Skala von 1 bis 4 bewertet, welche wiederum linear in eine Skala von 0 bis 100 umgewandelt wird. Der Wert 100 stellt dabei den schlechtesten Status dar (Chie et al. 2012). Verändert sich bei einem Patienten der Wert einer Skala um mindestens 10 Punkte im Vergleich zum Ausgangswert, so wird dieser Unterschied als klinisch relevant angesehen. Die Analysen des Endpunkts EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) wurden auf Basis der Daten der Patienten der ITT-Population durchgeführt, die einen Baseline- und mindestens einen Folgewert hatten. <u>Hauptanalyse:</u> Für die Analyse des Unterschieds zwischen Lenvatinib und Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items wurden das HR und das dazugehörige 95%-KI mithilfe des Cox-Proportional-Hazard-Modells für proportionale Hazards, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung“ als Kovariate, zum Signifikanzniveau α=0,05 berechnet. Zur Korrektur von Bindungen wurde die Efron-Methode verwendet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde außerdem mithilfe eines zweiseitigen stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, zum Signifikanzniveau α=0,05 analysiert. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das dazugehörige zweiseitige 95%-KI wurden mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. Die kumulativen Wahrscheinlichkeiten für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des jeweiligen Items wurden im Zeitverlauf mithilfe der Kaplan-Meier-Kurve dargestellt. <u>Sensitivitätsanalyse:</u> Für die Analyse der mittleren Veränderung zur Baseline im Score des jeweiligen Items wurden der Mittelwert zu Baseline und der LS-Mittelwert der Änderung zur Baseline für beide Behandlungsarme berechnet. Der Unterschied zwischen Lenvatinib und Sorafenib wurde mittels der Differenz der LS-Mittelwerte der Änderung zur Baseline analysiert. Bei der Analyse der mittleren Veränderung zur Baseline im Score des jeweiligen Items handelt es sich um eine Sensitivitätsanalyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; LS: Least Square

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-47: Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials für den Endpunkt EORTC QLQ-HCC18 beschrieben.

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95% (Tabelle 4-48). Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-HCC18 als hoch eingestuft (Tabelle 4-47).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden werden die Rücklaufquoten und die Ergebnisse der Haupt- und Sensitivitätsanalyse des Fragebogens EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) für die Studie 304 gemäß den in Abschnitt 4.2.5.2 erläuterten statistischen Verfahren dargestellt.

Rücklaufquoten

Tabelle 4-48: Ergebnisse der Bestimmung der Anteile auswertbarer Patienten für die Analyse des Fragebogens EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**n (%)**	**N**	**n (%)**
Körperbild	478	443 (92,7)	476	449 (94,3)
Ernährung	478	443 (92,7)	476	449 (94,3)
Sexualleben	478	434 (90,8)	476	429 (90,1)
EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); ITT: Intention-to-treat; n: Anzahl der Patienten in der ITT-Population mit Baseline- und mind. einem Folgewert; N: Anzahl der Patienten in der ITT-Population Quelle: SCO:SSiS 2018				

Die Rücklaufquoten des Fragebogens EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) sind in Tabelle 4-48 dargestellt. Hierbei handelt es sich um den Anteil derjenigen Patienten in der ITT-Population mit einem Baseline-Wert und mindestens einem Folgewert, die mit Hilfe des Fragebogens EORTC QLQ-HCC18 erfasst wurden.

Tabelle 4-49: Ergebnisse der Baseline-Werte gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**	
Endpunkt	**N**	**Mittelwert (SD) (Punkte)**	**N**	**Mittelwert (SD) (Punkte)**
Körperbild	463	15,1 (18,82)	459	14,2 (18,54)
Ernährung	463	14,7 (15,21)	459	12,3 (14,26)
Sexualleben	454	22,6 (31,60)	445	21,9 (31,26)
EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); N: Anzahl der Patienten zu Baseline; SD: Standardabweichung Quelle: SCO:SSiS 2018				

Die mittleren Baseline-Werte der Items gemessen auf den Skalen des Fragebogens EORTC QLQ-HCC18 sind in Tabelle 4-49 aufgelistet. Die Mittelwerte sowie die Standardabweichungen der beiden Therapiearme Lenvatinib und Sorafenib sind vergleichbar und lassen auf eine gute Randomisierung schließen.

Hauptanalyse

Im Folgenden werden die Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Items zur gesundheitsbezogenen Lebensqualität, gemessen auf den Skalen des Fragebogens EORTC QLQ-HCC18, dargestellt. Diese werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

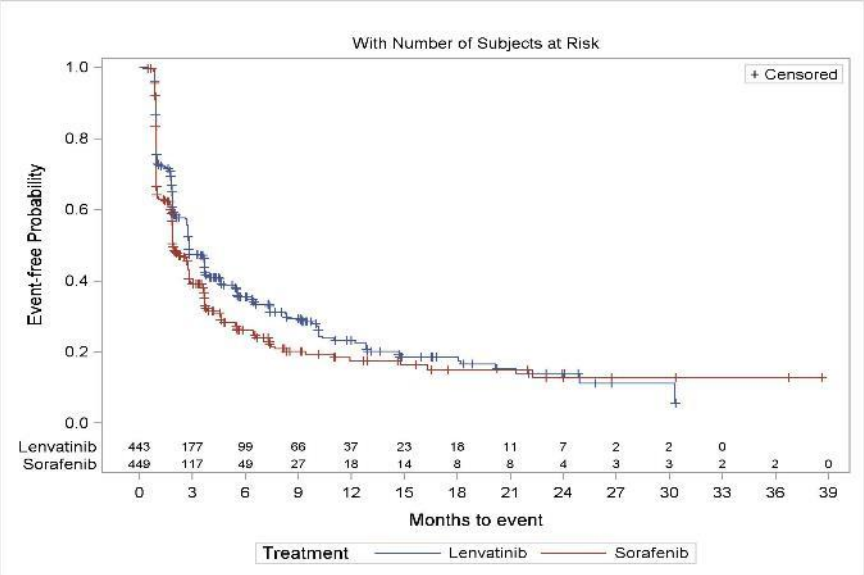


Abbildung 4-50: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

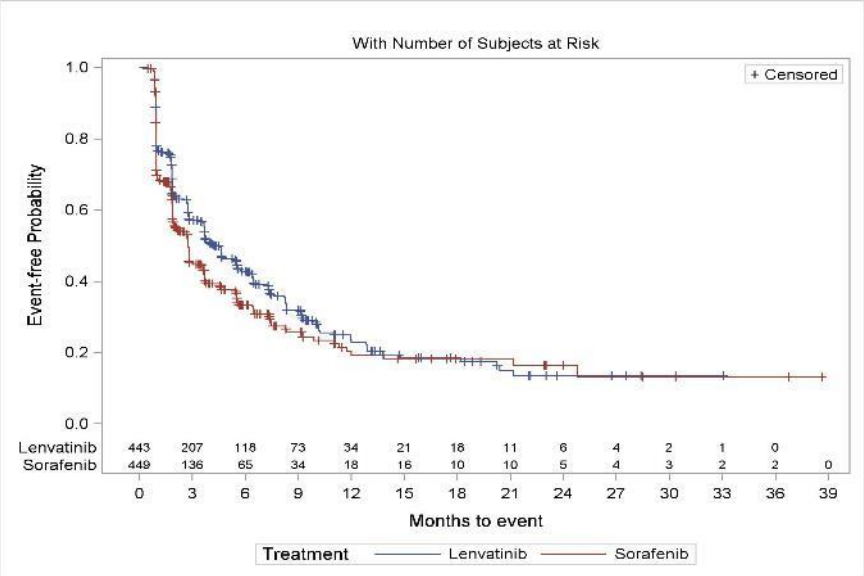


Abbildung 4-51: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

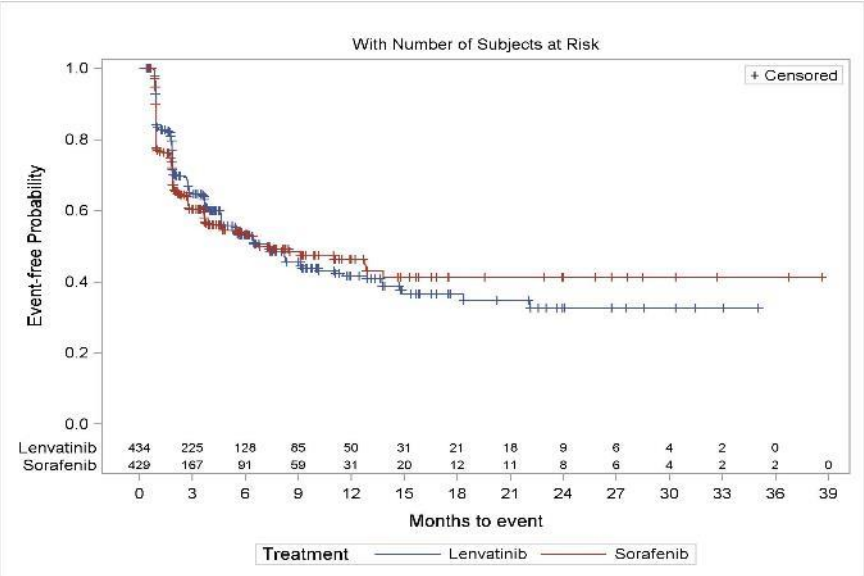


Abbildung 4-52: Hauptanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

Tabelle 4-50: Hauptanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib[c]**
Endpunkt	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**n/N (%)**	**Median[a] (Monate) [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Körperbild	299/443 (67,5)	2,8 [2,73; 3,68]	297/449 (66,1)	1,9 [1,87; 2,73]	0,79 [0,671; 0,932] 0,0051
Ernährung	279/443 (63,0)	4,3 [3,68; 5,52]	274/449 (61,0)	2,8 [2,07; 3,02]	0,80 [0,674; 0,949] 0,0102
Sexualleben	206/434 (47,5)	7,4 [5,49; 9,20]	183/429 (42,7)	6,7 [4,63; 13,80]	0,94 [0,771; 1,154] 0,5700

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurde anhand eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit „Behandlung" als Kovariate, berechnet.

c: Der p-Wert wurde anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, berechnet.

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire ; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; ITT: Intention-to-treat; IxRS: Interaktives Voice and Web Response Systems; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert

Quelle: SCO:SSiS 2018

Die Ergebnisse der Hauptanalyse des Fragebogens EORTC QLQ-HCC18 zeigte für die Items Körperbild und Ernährung eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell zeigte Lenvatinib gegenüber Sorafenib bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Körperbild eine statistisch signifikante Überlegenheit (HR=0,79; 95%-KI [0,671; 0,932]; p=0,0051). Eine statistisch signifikante Überlegenheit konnte auch bezüglich des Items Ernährung gezeigt werden (HR=0,80; 95%-KI [0,674; 0,949]; p=0,0102) (Tabelle 4-50).

Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte für das Item Körperbild betrug dabei 2,8 Monate (95%-KI [2,73; 3,68]) für Lenvatinib im Vergleich zu 1,9 Monaten (95%-KI [1,87; 2,73]) für Sorafenib. Bezüglich des Items Ernährung betrug der absolute Vorteil 1,5 Monate zugunsten von Lenvatinib (4,3 Monate; 95%-KI [3,68; 5,52]) im Vergleich zu Sorafenib (2,8 Monate; 95%-KI [2,07; 3,02]) (Tabelle 4-50).

Für das Item Sexualleben konnte ein numerischer Vorteil für Lenvatinib in der Zeit bis zur ersten Verschlechterung um 10 Punkte im Score des Items beobachtet werden (Tabelle 4-50).

Insgesamt belegen die Ergebnisse der Hauptanalyse eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Items Körperbild und Ernährung. Hinsichtlich des Items Sexualleben zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Sensitivitätsanalyse

Im Folgenden werden die Ergebnisse für die mittlere Änderung zur Baseline im Score der Items, gemessen auf den Skalen zur gesundheitsbezogenen Lebensqualität des EORTC QLQ-HCC18 (in Punkten), dargestellt. Diese werden im vorliegenden Dossier nicht für die Ableitung des Zusatznutzens herangezogen.

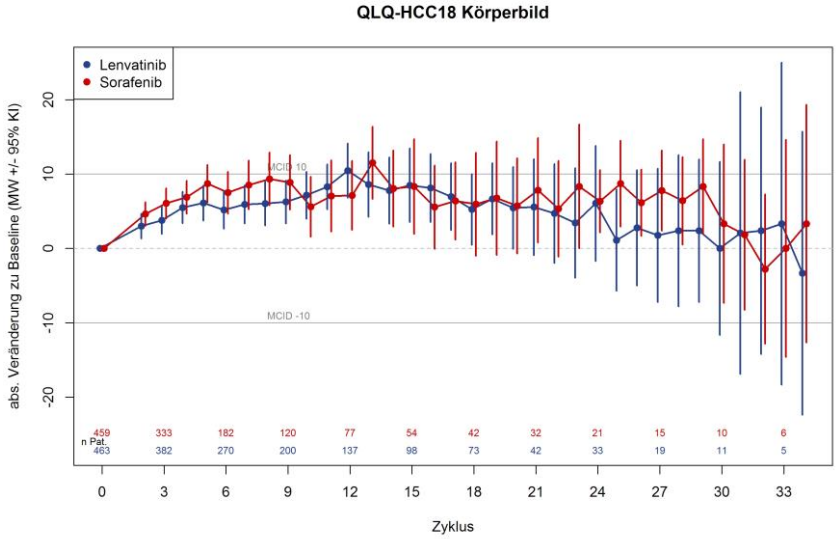


Abbildung 4-53: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

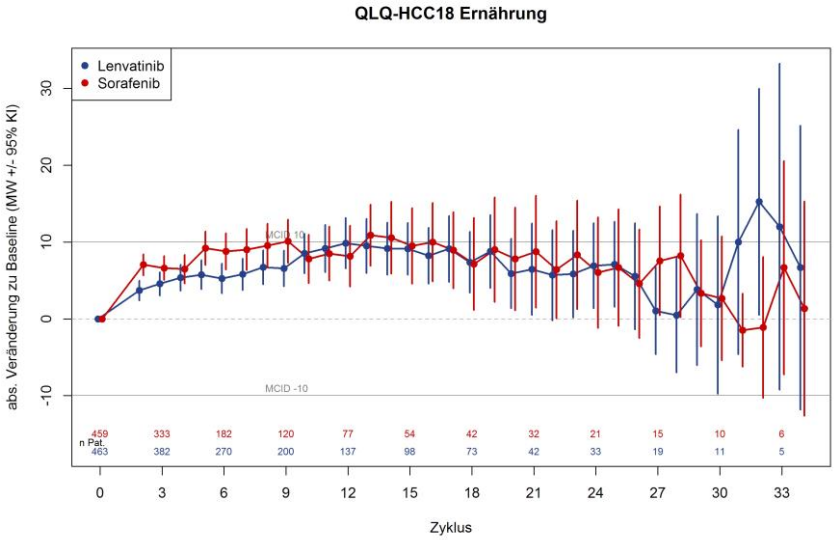


Abbildung 4-54: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

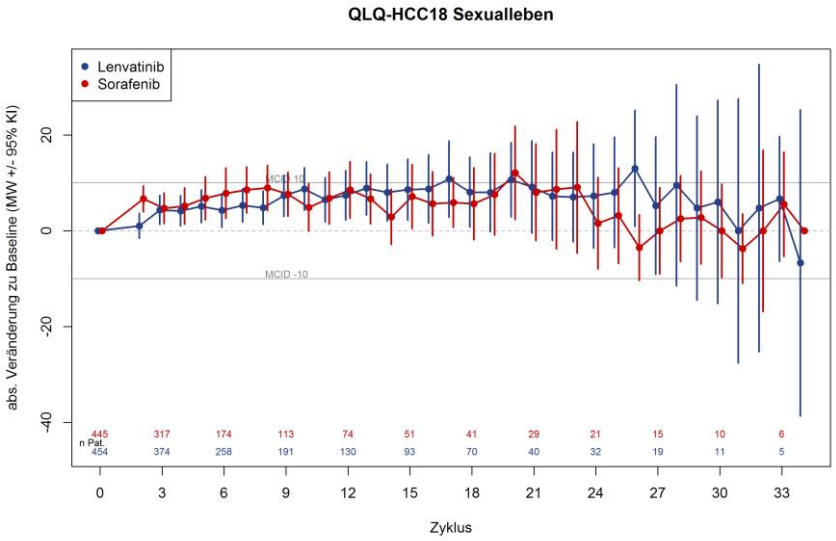


Abbildung 4-55: Sensitivitätsanalyse: Zeitverlauf der mittleren Änderung zur Baseline im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire

Quelle: SCO:SSiS 2018

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.4 Endpunkte zu Sicherheit und Verträglichkeit - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-51: Operationalisierung der Endpunkte zu Sicherheit und Verträglichkeit

Studie	Operationalisierung
304	Im vorliegenden Dossier erfolgt eine Darstellung und Untersuchung der folgenden Endpunkte zu Sicherheit und Verträglichkeit: • Unerwünschte/s Ereignisse (UE) • Schwerwiegende/s UE (SUE) • UE, die zum Therapieabbruch führten • UE mit CTCAE Grad ≥ 3 UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, werden ergänzend dargestellt. Die Kodierung der UE erfolgte unter Verwendung des MedDRA Version 19.1. Die Einteilung in verschiedene Grade erfolgte nach CTCAE Version 4.0. Alle UE und SUE wurden als TEAE erfasst. TEAE wurden definiert als jedes Ereignis, • das zu Baseline nicht präsent war und während oder bis zu 30 Tage nach der Behandlung mit der Studienmedikation auftrat. • das bei einer Vorbehandlung aufgetreten ist, aber zu Baseline nicht mehr präsent war und während oder bis zu 30 Tage nach der Behandlung mit der Studienmedikation wieder auftrat. • dessen Schweregrad sich während oder bis zu 30 Tage nach der Behandlung im Vergleich zum Zustand vor der Behandlung mit der Studienmedikation verschlechterte. Die Kodierung der UE erfolgte unter Verwendung des MedDRA Version 19.1. Die Einteilung in verschiedene Grade erfolgte nach Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. Die Analysen zu Sicherheit und Verträglichkeit wurden auf Basis der Daten der SAS-Population durchgeführt.

Hauptanalyse:

Für die Endpunkte zu Sicherheit und Verträglichkeit wurden Time-To-First-Event (Zeit bis zum ersten Auftreten) - Analysen durchgeführt. Das HR und das dazugehörige 95%-KI wurden mithilfe des unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

Die mediane Zeit bis zum ersten Ereignis wurde mittels der Kaplan-Meier-Methode für beide Behandlungsarme kalkuliert. Das zur medianen Zeit gehörige zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

Der Zusatznutzen wird aufgrund der unterschiedlichen Gesamtbeobachtungszeiten und der damit einhergehenden Verzerrung der Inzidenzraten ausschließlich auf Basis der Time-To-First-Event-Analysen mittels HR und 95%-KI hergeleitet.

Ergänzende Analyse:

Ergänzend zu den kategorialen Verträglichkeitsendpunkten werden zur Darstellung des Verträglichkeitsprofils die Analyse der Einzel-UE mit CTCAE Grad 3 oder 4 betrachtet, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag.

- Diarrhö mit CTCAE Grad 3 oder 4
- Erhöhte Gamma-Glutamyl Transferase mit CTCAE Grad 3 oder 4
- Hepatische Enzephalopathie mit CTCAE Grad 3 + 4
- Fatigue mit CTCAE Grad 3 oder 4
- Aszites mit CTCAE Grad 3 oder 4
- Hyponatriämie mit CTCAE Grad 3 oder 4
- Anämie mit CTCAE Grad 3 oder 4
- Erhöhte Alaninaminotransferase mit CTCAE Grad 3 oder 4
- Palmar-plantares Erythrodysästhesiesyndrom mit CTCAE Grad 3 oder 4
- Asthenie mit CTCAE Grad 3 oder 4
- Erniedrigte Leukozytenzahl mit CTCAE Grad 3 oder 4

Bei der Analyse der Einzel-UE mit CTCAE Grad 3 oder 4 handelt es sich um eine ergänzende Analyse, deren Ergebnisse nicht für die Ableitung des Zusatznutzens herangezogen werden.

CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; SAS: Safety Analysis Set; SUE: Schwerwiegendes unerwünschtes Ereignis / schwerwiegende unerwünschte Ereignisse; TEAE: Treatment Emergent Adverse Event; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-52: Bewertung des Verzerrungspotenzials für die Endpunkte zu Sicherheit und Verträglichkeit[a] in Studie 304 mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
304	niedrig	nein	ja	ja	ja	niedrig

a: Das Verzerrungspotential der Endpunkte Anzahl Patienten mit UE, SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 wurde identisch bewertet und wird daher hier zusammengefasst dargestellt

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie 304 wurde als Open-Label-Studie durchgeführt, deren Verzerrungspotenzial auf Studienebene als niedrig einzustufen ist. Im Folgenden wird die Bewertung des Verzerrungspotenzials der Endpunkte zu Sicherheit und Verträglichkeit beschrieben.

Das Verzerrungspotential der Endpunkte zu Sicherheit und Verträglichkeit wird als niedrig eingestuft, da eine Beeinflussung hinsichtlich des Ergebnisses der Endpunkte auch bei einem Open-Label-Studiendesign durch das Wissen um die Behandlung weder durch den Patienten noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis des SAS, in das alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhalten haben. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor (Tabelle 4-52).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Analysen

Im Folgenden sind die Ergebnisse der Analyse zur Sicherheit und Verträglichkeit der Studie 304 für die SAS-Population dargestellt. Vor der Darstellung der Ergebnisse erfolgt zunächst eine Übersicht der Gesamtbeobachtungszeiten der UE in der Studie 304. Anschließend werden die Ergebnisse der Analysen der Zeit bis zum Auftreten des ersten UE (in Tagen) und die ergänzenden Analysen tabellarisch dargestellt.

Übersicht der Gesamtbeobachtungszeiten der UE

Tabelle 4-53: Zusammenfassung der Länge der Gesamtbeobachtungszeit der UE (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel

Länge der Gesamtbeobachtungszeit der UE (Tage)[a]	**Lenvatinib**	**Sorafenib**
N	476	475
Mittelwert (SD)	265,7 (211,93)	201,3 (195,49)
Median (Minimum; Maximum)	200,5 (1; 1074)	125,0 (13; 1176)
a: Länge der Gesamtbeobachtungszeit der UE = Letztes Datum der Auswertung der Sicherheitsanalyse - Datum der ersten Dosierung des Studienpräparats + 1 N: Anzahl der Patienten in der SAS-Population; SAS: Safety Analysis Set; SD: Standard Deviation (Standardabweichung); UE: Unerwünschtes Ereignis / unerwünschte Ereignisse Quelle: Eisai Inc. 2018c		

Wie in Tabelle 4-53 dargestellt, war die mediane Gesamtbeobachtungszeit für Lenvatinib um mehr als 75 Tage länger als für Sorafenib. Durch die deutlich längeren Gesamtbeobachtungszeiten wären die Inzidenzraten und die darauf basierenden Effektmaße RR, OR und RD zuungunsten von Lenvatinib verzerrt, so dass nur sehr eingeschränkt Aussagen zu einem größeren oder vergleichbaren bzw. geringeren Schaden auf Basis der Inzidenzraten gemacht werden könnten. Deswegen basiert die Bewertung der Sicherheit und Verträglichkeit von Lenvatinib im vorliegenden Dossier alleine auf den Analysen der Zeit bis zum Auftreten des ersten Ereignisses (Time-To-First-Event-Analysen).

Hauptanalyse

Im Folgenden werden die Ergebnisse für die Zeit bis zum Auftreten des ersten UE (in Tagen) dargestellt. Diese werden im vorliegenden Dossier für die Ableitung des Zusatznutzens herangezogen.

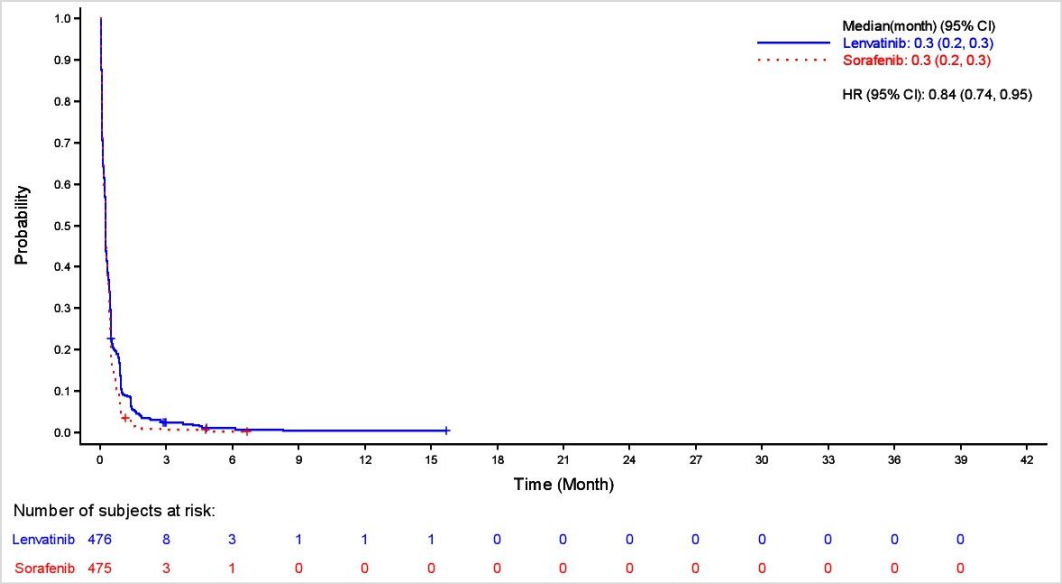


Abbildung 4-56: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; UE: Unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

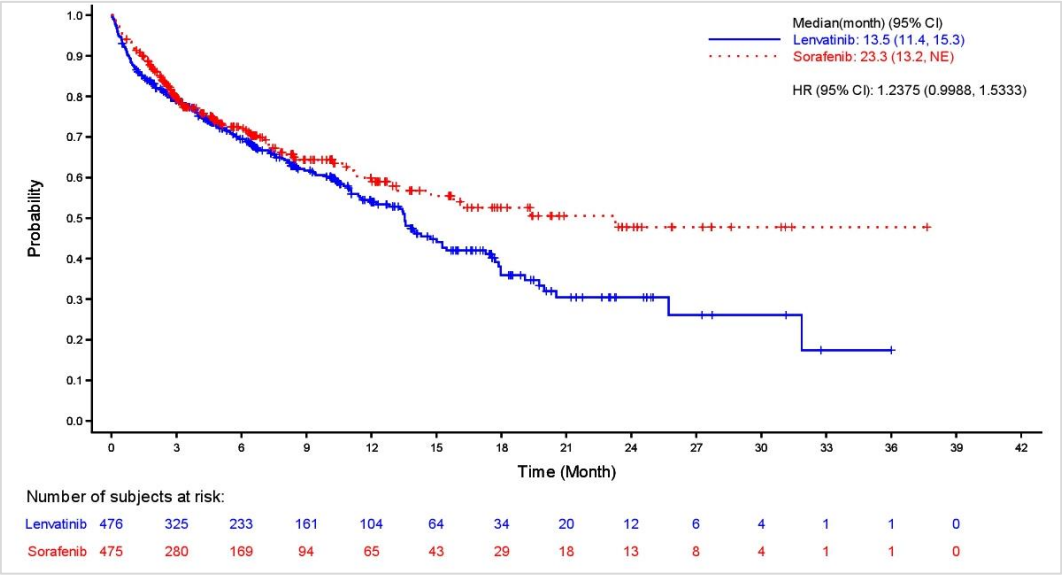


Abbildung 4-57: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **SUE** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; SUE: Schwerwiegende unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

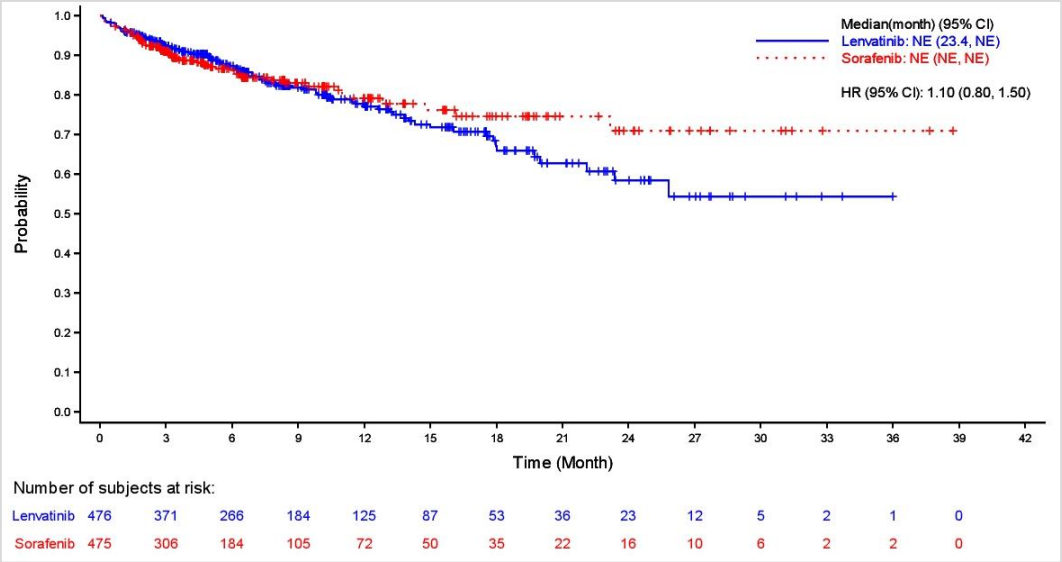


Abbildung 4-58: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE, die zum Therapieabbruch führten** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; UE: Unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

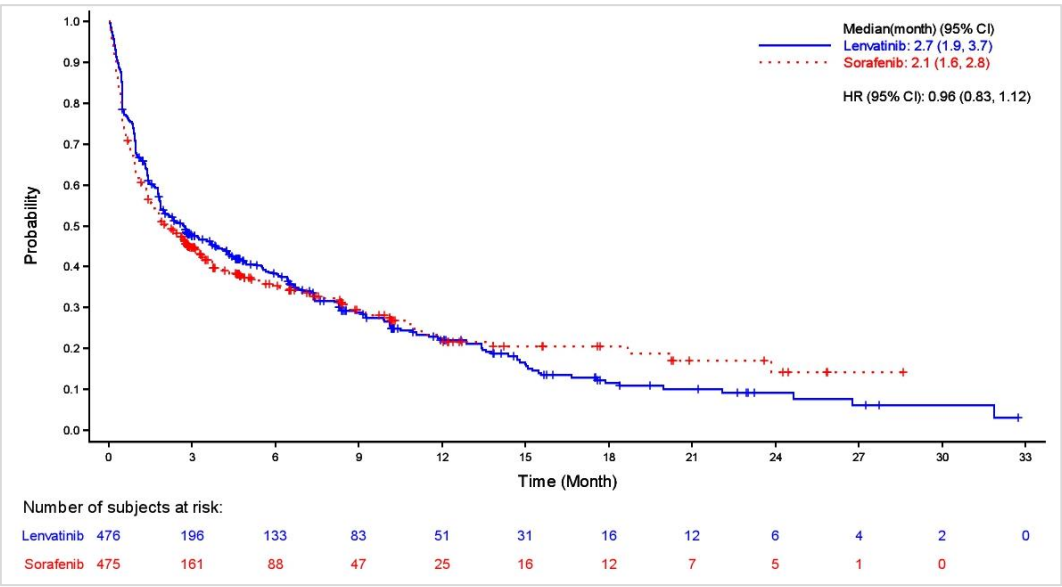


Abbildung 4-59: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; UE: Unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

Tabelle 4-54: Hauptanalyse: Ergebnisse für die Zeit bis zum Auftreten des ersten UE (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel

Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
n/N (%)	**Median[a] (Tage) [95% KI]**	**n/N (%)**	**Median[a] (Tage) [95% KI]**	**HR[b] [95% KI]**
Patienten mit UE				
470/476 (98,7)	8 [7; 8]	472/475 (99,4)	8 [7; 8]	0,84 [0,74; 0,95]
Patienten mit SUE				
205/476 (43,1)	412 [347; 465]	144/475 (30,3)	708 [402; NE]	1,24 [0,999; 1,53]
Patienten mit UE, die zum Therapieabbruch führten				
94/476 (19,8)	NE [711; NE]	69/475 (14,5)	NE [NE; NE]	1,10 [0,80; 1,50]
Patienten mit UE mit CTCAE Grad ≥ 3				
357/476 (75,0)	82 [57; 112]	316/475 (66,5)	64 [50; 86]	0,96 [0,83; 1,12]

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurden mithilfe des unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der SASPopulation; NE: not estimable (nicht berechenbar); SAS: Safety Analysis Set; SUE: Schwerwiegendes unerwünschtes Ereignis / schwerwiegende unerwünschte Ereignisse; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

Für die Time-To-First-Event-Analyse ergaben sich für die Kategorien SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 keine statistisch signifikanten Unterschiede zwischen den Studienarmen Lenvatinib und Sorafenib. Bezüglich der Kategorie UE zeigte sich bei der Analyse eine Überlegenheit für Lenvatinib im Vergleich zu Sorafenib (HR=0,84; 95%-KI [0,74; 0,95]).

Ergänzende Analyse

Ergänzend zu den dargestellten Ergebnissen der Hauptanalyse der Studie 304 sind im Folgenden die Ergebnisse für die UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, dargestellt. Die Ergebnisse werden nicht zur Ableitung des Zusatznutzens herangezogen.

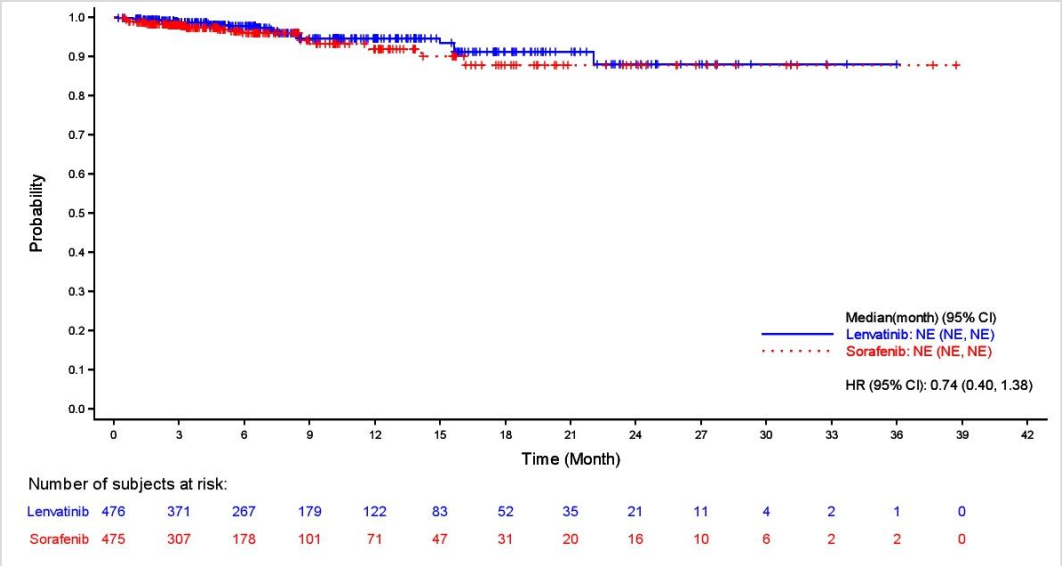


Abbildung 4-60: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Diarrhö mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

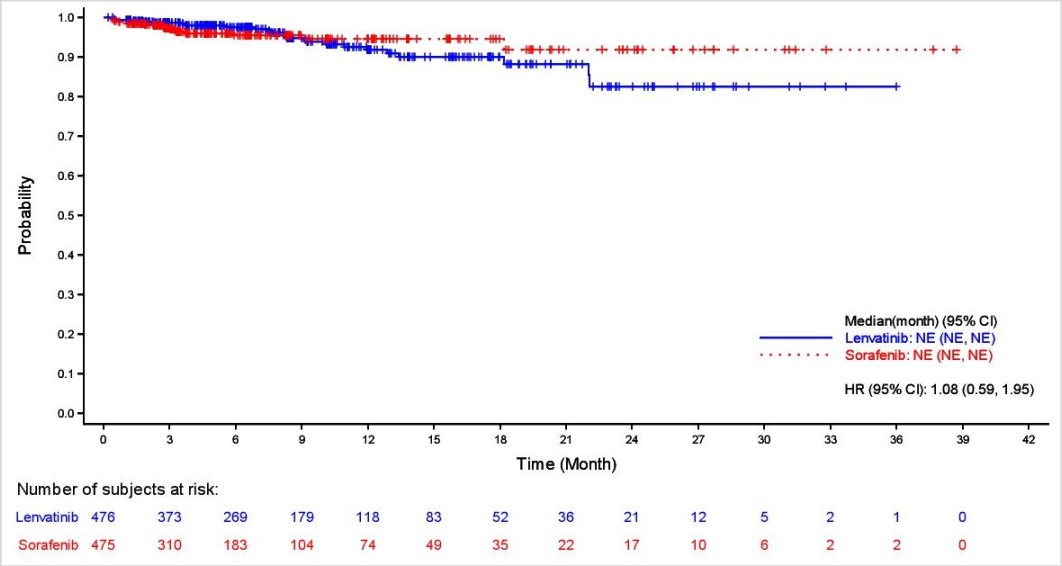


Abbildung 4-61: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte Gamma-Glutamyl Transferase mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

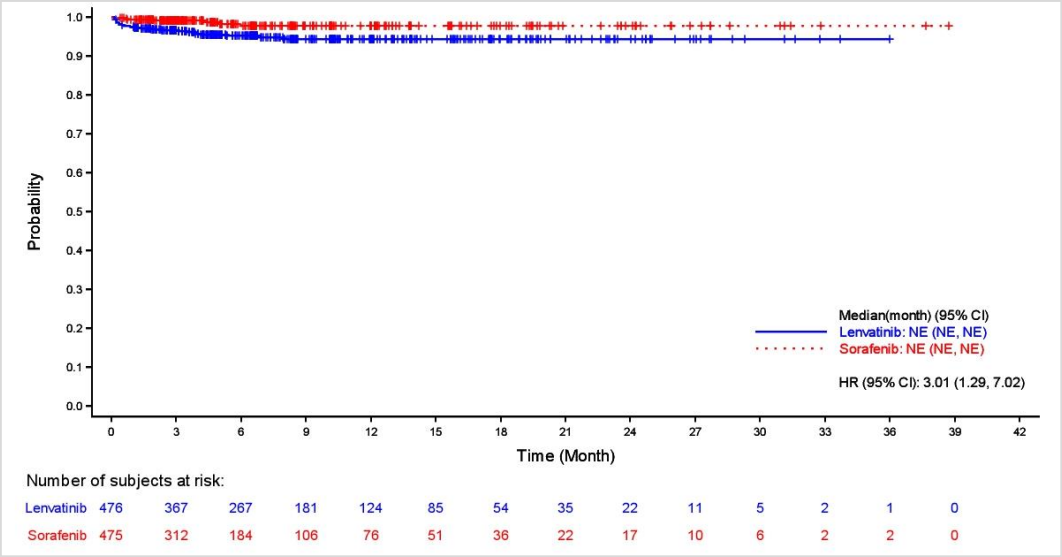


Abbildung 4-62: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **hepatische Enzephalopathie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

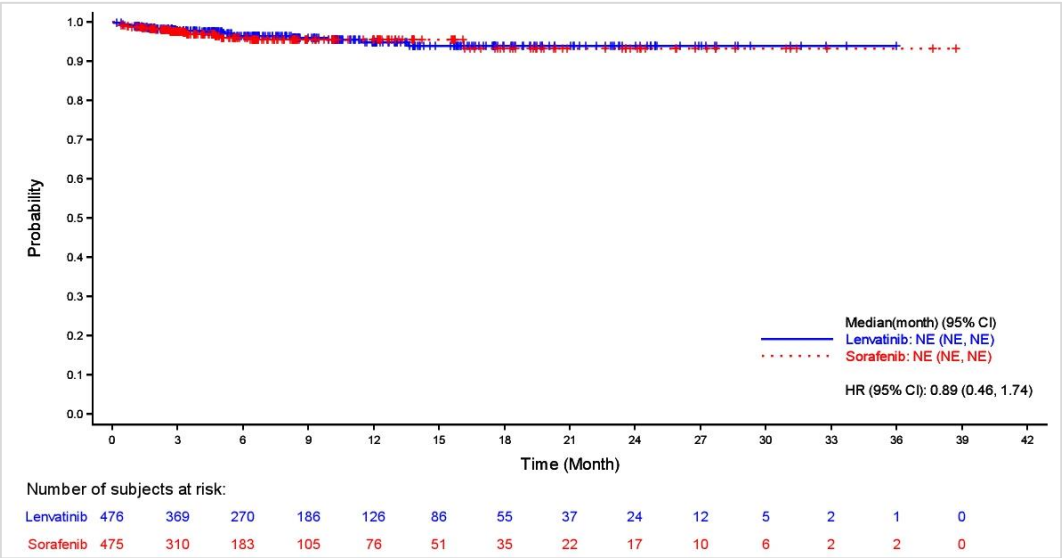


Abbildung 4-63: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Fatigue mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

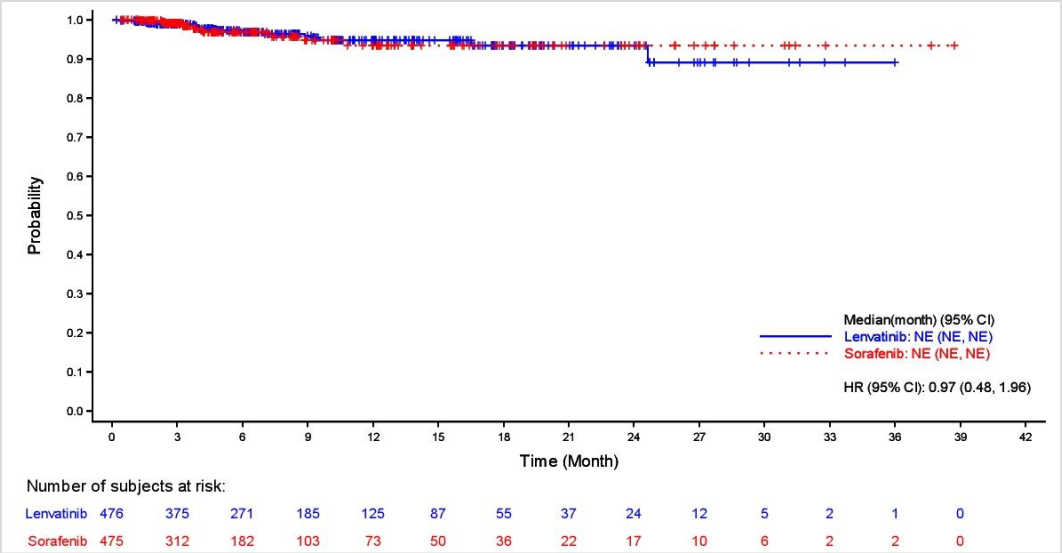


Abbildung 4-64: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Aszites mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio

Quelle: Eisai Inc. 2018c

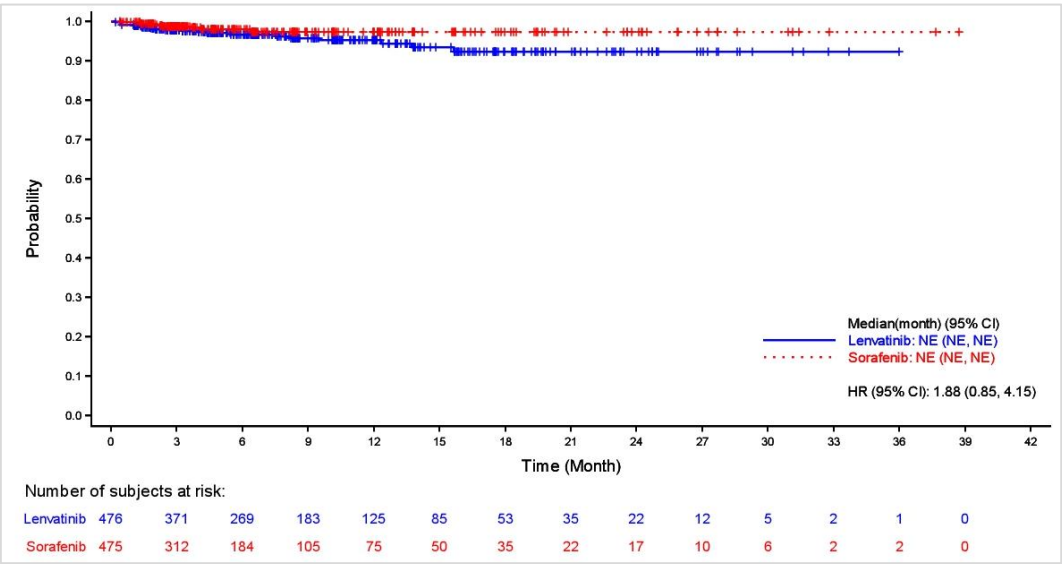


Abbildung 4-65: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Hyponatriämie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio

Quelle: Eisai Inc. 2018c

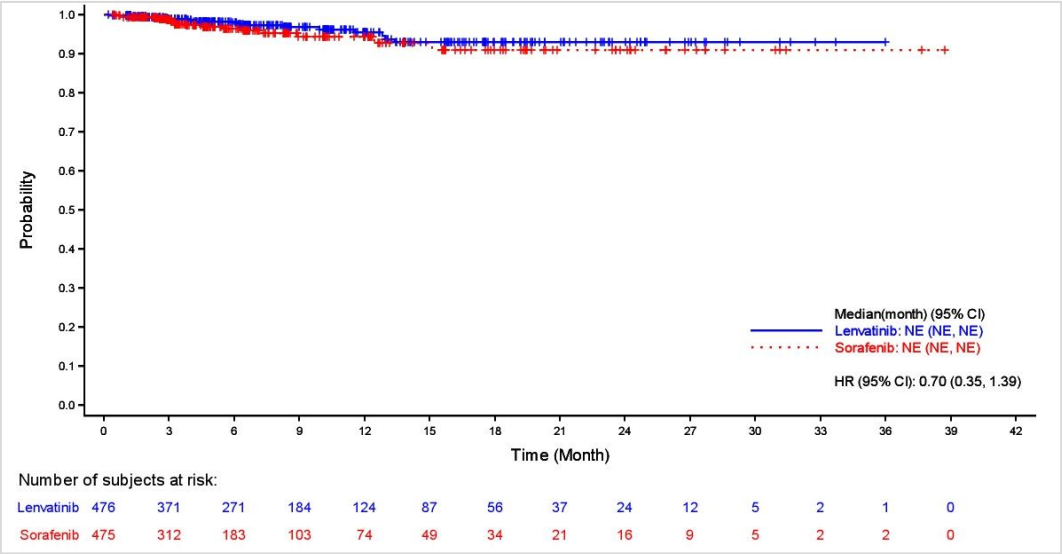


Abbildung 4-66: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Anämie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

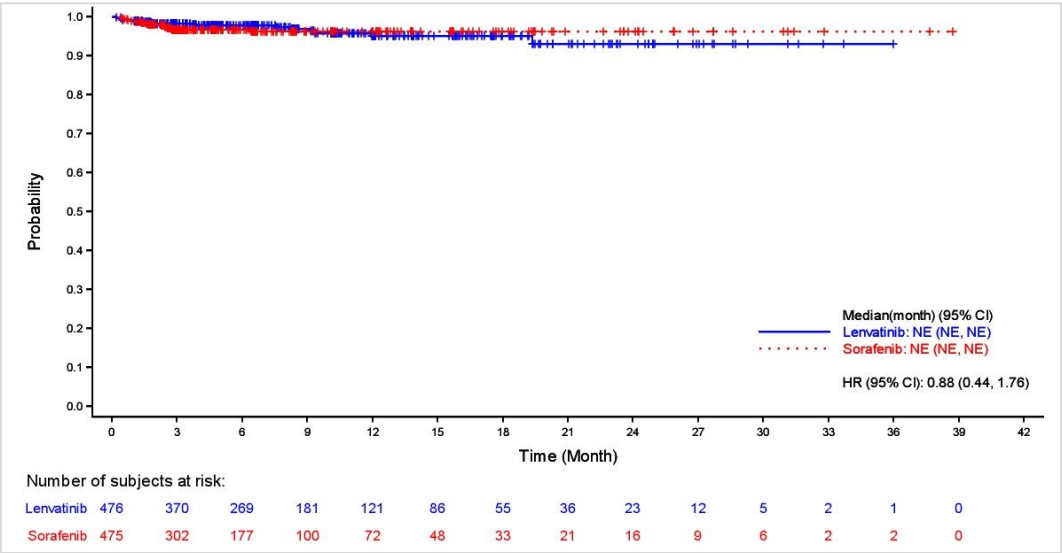


Abbildung 4-67: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte Alaninaminotransferase mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

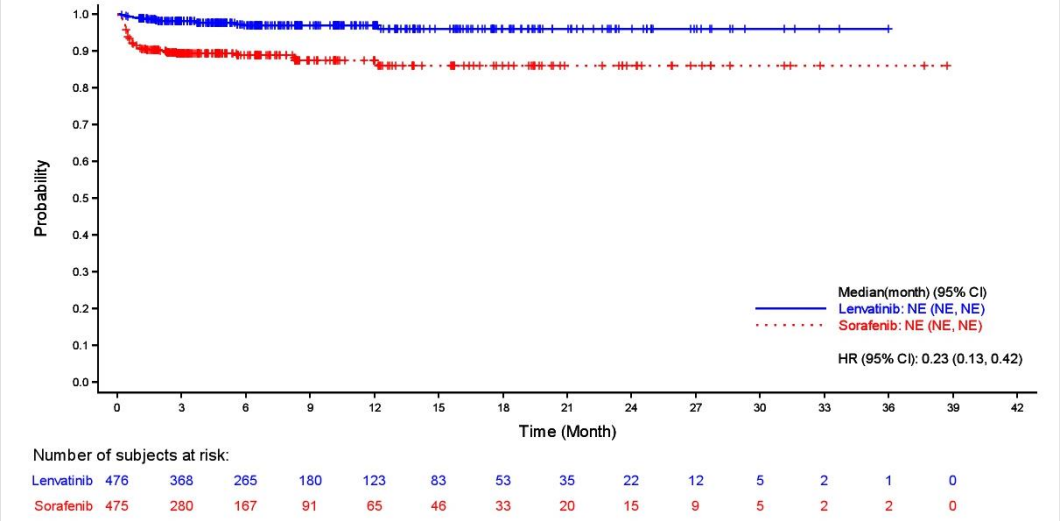


Abbildung 4-68: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **palmar-plantares Erythrodysästhesiesyndrom mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

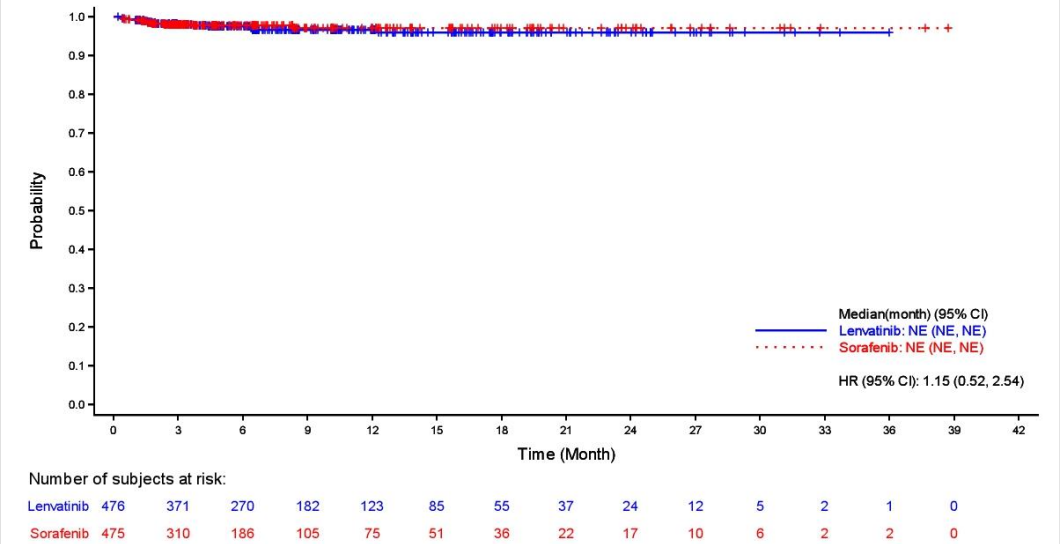


Abbildung 4-69: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Asthenie mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

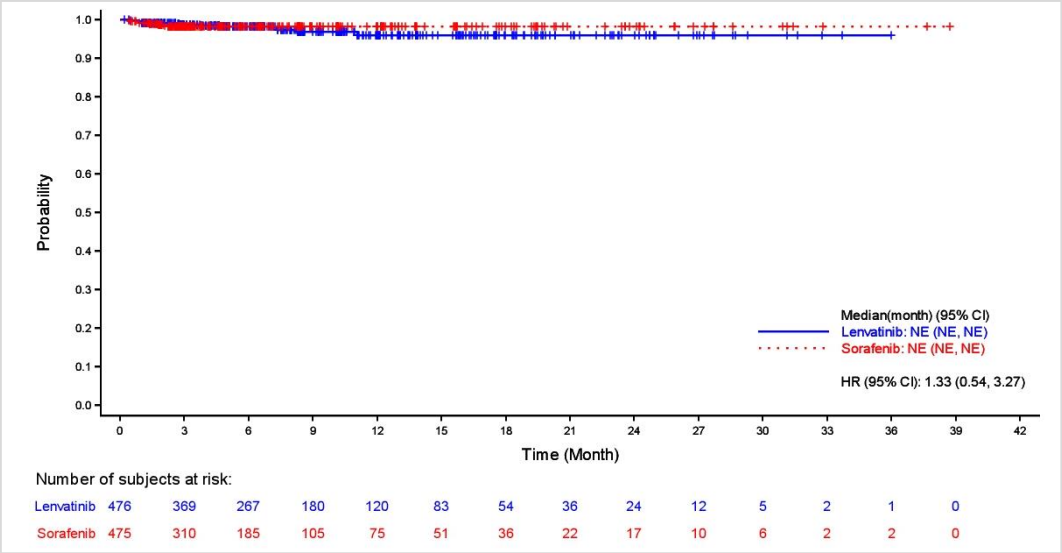


Abbildung 4-70: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erniedrigte Leukozytenzahl mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

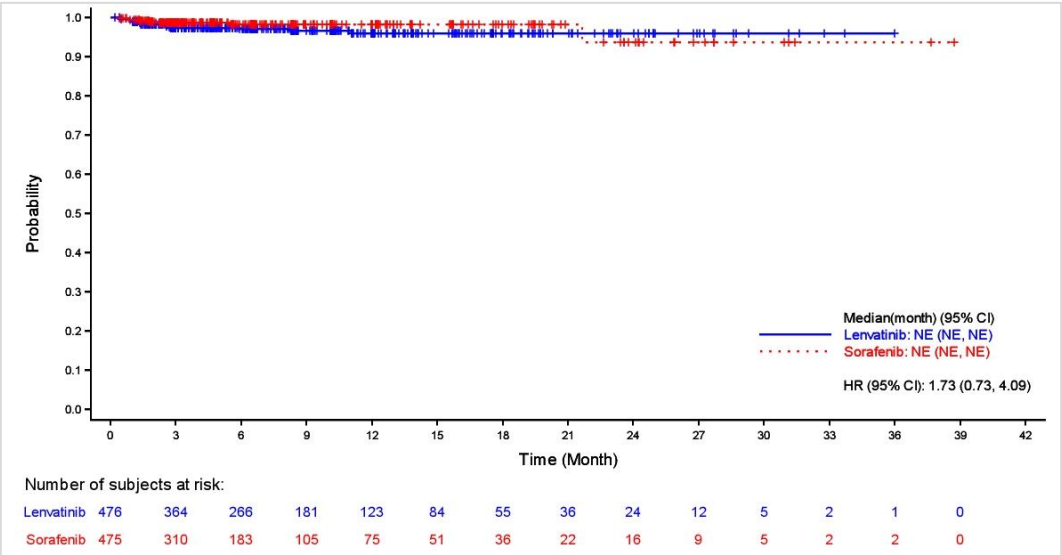


Abbildung 4-71: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erniedrigte Neutrophilenzahl mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

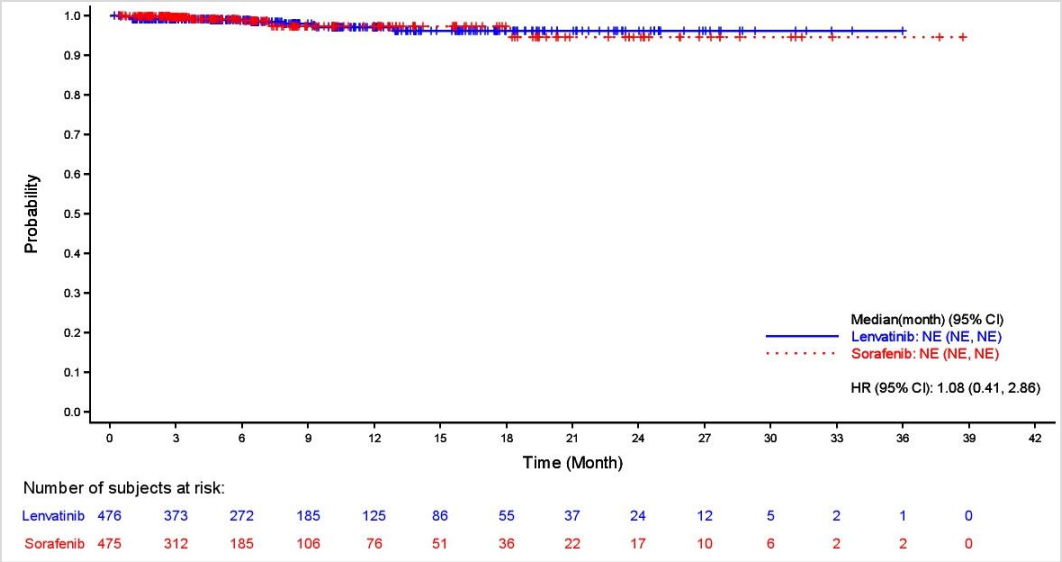


Abbildung 4-72: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte alkalische Phosphatase im Blut mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

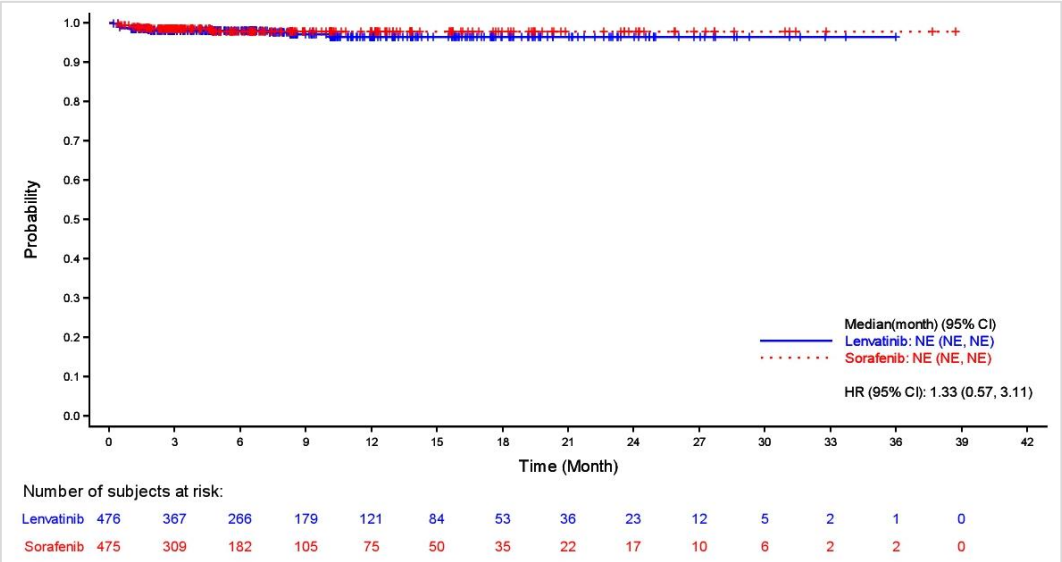


Abbildung 4-73: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **erhöhte Lipase mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

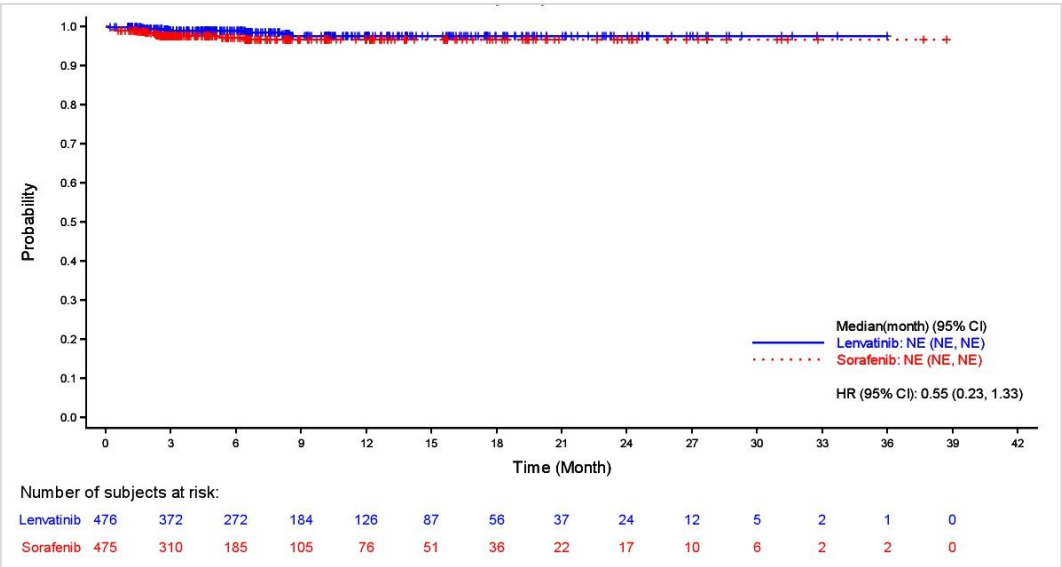


Abbildung 4-74: Hauptanalyse: Kaplan-Meier-Kurve für den Endpunkt **Abdominalschmerz mit CTCAE Grad 3 oder 4** aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio
Quelle: Eisai Inc. 2018c

Tabelle 4-55: Ergänzende Analyse: Ergebnisse für die Zeit bis zum Auftreten des ersten UE mit CTCAE Grad 3 oder 4 (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel

Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
n/N (%)	**Median[a] (Tage) [95% KI]**	**n/N (%)**	**Median[a] (Tage) [95% KI]**	**HR[b] [95% KI]**
Patienten mit Diarrhö mit CTCAE Grad 3 oder 4				
20/476 (4,2)	NE [NE; NE]	20/475 (4,2)	NE [NE; NE]	0,74 [0,40; 1,38]
Patienten mit erhöhter Gamma-Glutamyl Transferase mit CTCAE Grad 3 oder 4				
26/476 (5,5)	NE [NE; NE]	19/475 (4,0)	NE [NE; NE]	1,08 [0,59; 1,95]
Patienten mit hepatischer Enzephalopathie mit CTCAE Grad 3 oder 4				
23/476 (4,8)	NE [NE; NE]	7/475 (1,5)	NE [NE; NE]	3,01 [1,29; 7,02]
Patienten mit Fatigue mit CTCAE Grad 3 oder 4				
18/476 (3,8)	NE [NE; NE]	17/475 (3,6)	NE [NE; NE]	0,89 [0,46; 1,74]
Patienten mit Aszites mit CTCAE Grad 3 oder 4				
18/476 (3,8)	NE [NE; NE]	14/475 (3,0)	NE [NE; NE]	0,97 [0,48; 1,96]

Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
n/N (%)	**Median[a] (Tage) [95% KI]**	**n/N (%)**	**Median[a] (Tage) [95% KI]**	**HR[b] [95% KI]**
Patienten mit Hyponatriämie mit CTCAE Grad 3 oder 4				
20/476 (4,2)	NE [NE; NE]	9/475 (1,9)	NE [NE; NE]	1,88 [0,85; 4,15]
Patienten mit Anämie mit CTCAE Grad 3 oder 4				
16/476 (3,4)	NE [NE; NE]	17/475 (3,6)	NE [NE; NE]	0,70 [0,35; 1,39]
Patienten mit erhöhter Alaninaminotransferase mit CTCAE Grad 3 oder 4				
16/476 (3,4)	NE [NE; NE]	16/475 (3,4)	NE [NE; NE]	0,88 [0,44; 1,76]
Patienten mit palmar-plantarem Erythrodysästhesiesyndrom mit CTCAE Grad 3 oder 4				
14/476 (2,9)	NE [NE; NE]	54/475 (11,4)	NE [NE; NE]	0,23 [0,13; 0,42]
Patienten mit Asthenie mit CTCAE Grad 3 oder 4				
14/476 (2,9)	NE [NE; NE]	11/475 (2,3)	NE [NE; NE]	1,15 [0,52; 2,54]
Patienten mit erniedrigter Leukozytenzahl mit CTCAE Grad 3 oder 4				
12/476 (2,5)	NE [NE; NE]	8/475 (1,7)	NE [NE; NE]	1,33 [0,54; 3,27]
Patienten mit erniedrigter Neutrophilenzahl mit CTCAE Grad 3 oder 4				
15/476 (3,2)	NE [NE; NE]	8/475 (1,7)	NE [NE; NE]	1,73 [0,73; 4,09]
Patienten mit erhöhter alkalische Phosphatase im Blut mit CTCAE Grad 3 oder 4				
10/476 (2,1)	NE [NE; NE]	7/475 (1,5)	NE [NE; NE]	1,08 [0,41; 2,86]
Patienten mit erhöhter Lipase mit CTCAE Grad 3 oder 4				
13/476 (2,7)	NE [NE; NE]	9/475 (1,9)	NE [NE; NE]	1,33 [0,57; 3,11]
Patienten mit Abdominalschmerz mit CTCAE Grad 3 oder 4				
8/476 (1,7)	NE [NE; NE]	13/475 (2,7)	NE [NE; NE]	0,55 [0,23; 1,33]

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR mit dem dazugehörigen 95%-KI und p-Wert wurden mithilfe des unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; KI: Konfidenzintervall; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der SASPopulation; NE: not estimable (nicht berechenbar); SAS: Safety Analysis Set; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

Für die Time-To-First-Event-Analyse ergaben sich für die UE mit CTCAE Grad 3 oder 4, mit Ausnahme des UE hepatische Enzephalopathie und palmar-plantares Erythrodysästhesiesyndrom keine statistisch signifikanten Unterschiede zwischen den Studienarmen Lenvatinib und Sorafenib. Bezüglich der Zeit bis zum ersten UE hepatische Enzephalopathie mit CTCAE Grad 3 oder 4 ergab sich ein statistisch signifikanter Nachteil für Lenvatinib gegenüber Sorafenib (HR=3,01; 95%-KI [1,29; 7,02]). Bezüglich der Zeit bis zum ersten UE palmar-plantares Erythrodysästhesiesyndrom mit CTCAE Grad 3 oder 4 ergab sich ein statistisch signifikanter Vorteil für Lenvatinib gegenüber Sorafenib (HR=0,23; 95 5-KI [0,13; 0,42]).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Es wurde keine Meta-Analyse durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Zur Untersuchung der Effektmodifikation wurden a priori und post-hoc geplante Subgruppenanalysen für alle patientenrelevanten Endpunkte der Studie 304 durchgeführt (siehe Abschnitt 4.2.5).

Im Folgenden werden die Ergebnisse der Subgruppenanalysen und die entsprechenden Interaktions-p-Werte der Tests auf Interaktion tabellarisch dargestellt. Dabei wurden Interaktions-p-Werte < 0,05 als Belege für eine Interaktion interpretiert.

4.3.1.3.2.1 Subgruppenanalyse für den Endpunkt Mortalität

4.3.1.3.2.1.1 Subgruppenanalyse für den Endpunkt Gesamtüberleben (OS)

Tabelle 4-56: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für den Endpunkt OS aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	**Interaktions-p-Wert[a]**
Baseline Altersgruppe,	0,7989
Geschlecht	0,9284
Region	0,1335
MPVI	0,6784
ES	0,4324
MPVI oder ES oder beides	0,2952
Baseline ECOG-PS	0,4607
Baseline Gewichtsgruppe	0,4278
Baseline AFP	0,4434
Ätiologie HBV	0,2529
Ätiologie HCV	0,4929
Ätiologie Alkohol	0,8943
BCLC Staging	0,9931
Leberzirrhose zu Baseline (CRF)	0,7457
Leberzirrhose zu Baseline (IIR)	0,4139
Vorbehandlung mit TACE	0,3171

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); OS: Overall Survival (Gesamtüberleben); TACE: Transarterielle Chemoembolisation

Quelle: Eisai Inc. 2018c

Es wurde für OS kein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bzgl. der in Tabelle 4-56 aufgelisteten Subgruppenmerkmale festgestellt. Die Ergebnisse der Subgruppenanalysen werden aufgrund dessen hier nicht dargestellt, sind aber in Modul 5 zu finden.

4.3.1.3.2.2 Subgruppenanalyse für den Endpunkt Morbidität

4.3.1.3.2.2.1 Subgruppenanalyse für den Endpunkt progressionsfreies Überleben (PFS)

Tabelle 4-57: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für den Endpunkt PFS (IIR, mRECIST) aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	**Interaktions-p-Wert[a]**
Baseline Altersgruppe,	0,9390
Geschlecht	0,2166
Region	**0,0063**
MPVI	0,6197
ES	0,4298
MPVI oder ES oder beides	0,6615
Baseline ECOG-PS	0,7266
Baseline Gewichtsgruppe	0,4773
Baseline AFP	0,3355
Ätiologie HBV	0.5903
Ätiologie HCV	0.5820
Ätiologie Alkohol	0.9699
BCLC Staging	0,4379
Leberzirrhose zu Baseline (CRF)	0,8978
Leberzirrhose zu Baseline (IIR)	0,9787
Vorbehandlung mit TACE	0,2878

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); mRECIST: modified Response Evaluation Criteria in Solid Tumors; PFS: Progression-Free Survival (progressionsfreies Überleben); TACE: Transarterielle Chemoembolisation

Quelle: Eisai Inc. 2018c

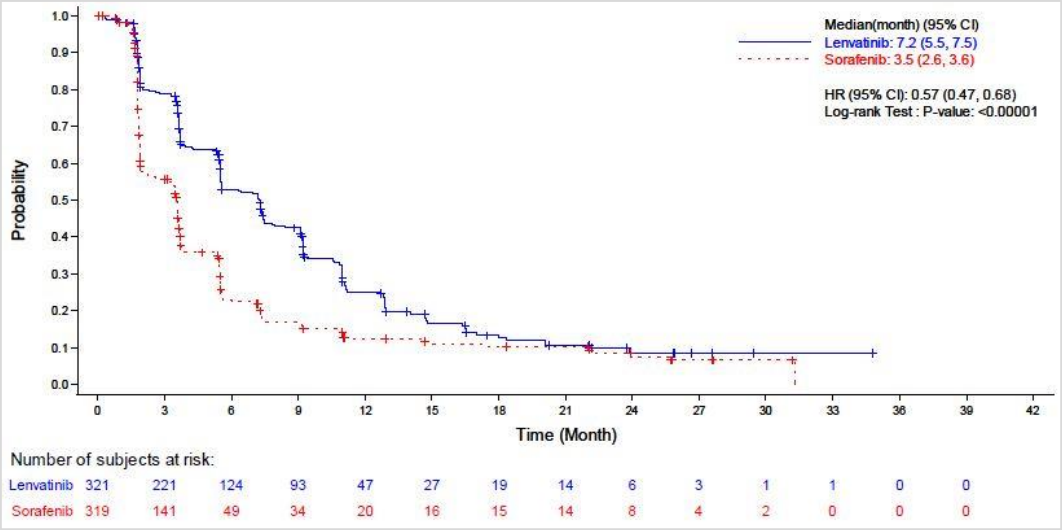


Abbildung 4-75: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **PFS** (IIR, mRECIST) (Region[a] „Asien-Pazifik") aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); mRECIST: modified Response Evaluation Criteria in Solid Tumors; PFS: Progression-Free Survival (progressionsfreies Überleben)
Quelle: Eisai Inc. 2018b

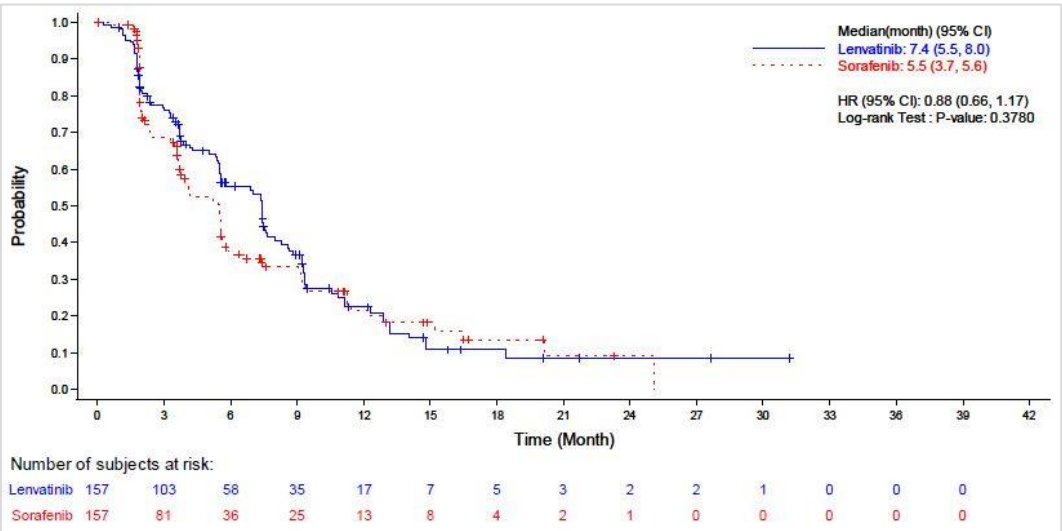


Abbildung 4-76: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt PFS (IIR, mRECIST) (Region[a] „Westliche Regionen") aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); mRECIST: modified Response Evaluation Criteria in Solid Tumors; PFS: Progression-Free Survival (progressionsfreies Überleben)
Quelle: Eisai Inc. 2018b

Tabelle 4-58: Subgruppenanalyse: Ergebnisse für den Endpunkt PFS (IIR, mRECIST) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Kategorie	**n/N (%)**	**Median[b] (Monate) [95% KI]**	**n/N (%)**	**Median[b] (Monate) [95% KI]**	**HR[c] [95% KI] p-Wert[d]**
Progressionsfreies Überleben (IIR, mRECIST)					
Alle	311/478 (65,1)	7,3 [5,6; 7,5]	323/476 (67,9)	3,6 [3,6; 3,7]	0,64 [0,55; 0,75] <0,0001
Region[a]					
Asien-Pazifik	212/321 (66,0)	7,2 [5,5; 7,5]	232/319 (72,7)	3,5 [2,6; 3,6]	0,57 [0,47; 0,68] 0,0001
Westliche Regionen	99/157 (63,1)	7,4 [5,5; 8,0]	91/157 (58,0)	5,5 [3,7; 5,6)	0,88 [0,66; 1,17] 0,3780

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.

b: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

c: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

d: Der p-Wert wurde anhand eines unstratifizierten Log-Rank-Tests berechnet.

HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); ITT: Intention-to-treat; KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population; PFS: Progression-Free Survival (progressionsfreies Überleben)

Quelle: Eisai Inc. 2018c

In der Studie 304 wurde für PFS ein Beleg für Interaktion (Interaktion-p-Wert < 0,05) bzgl. des Subgruppenmerkmals „Region“ festgestellt (Tabelle 4-57). Hinsichtlich dieses Merkmals ist der Effekt für Patienten der westlichen Regionen sowie aus der Region Asien-Pazifik gleichgerichtet wie in der Gesamtpopulation (Tabelle 4-58). Damit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.1.3.2.2.2 Subgruppenanalyse für den Endpunkt EORTC QLQ-C30 (Symptome)

Tabelle 4-59: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	Interaktions-p-Wert[a]							
	Fatigue	Übelkeit und Erbrechen	Schmerz	Dys-pnoe	Schlaf-losigkeit	Appetit-minderung	Obsti-pation	Diarrhö
Baseline Altersgruppe,	0,6011	0,8758	0,2793	0,3766	0,0576	0,3056	0,7637	**0,0451**
Geschlecht	0,3325	0,8027	0,8983	0,8352	0,3677	0,2705	0,0700	0,0945
Region	0,1272	0,8150	0,0984	0,2961	0,3835	0,9336	0,9192	0,3815
MPVI	0,7532	0,6532	0,6412	**0,0146**	**0,0075**	0,1563	0,4268	0,3289
ES	0,6956	0,0630	0,4604	0,6682	0,8567	0,5518	0,9913	0,6765
MPVI oder ES oder beides	0,1096	**0,0067**	0,4938	0,3286	0,2785	0,1004	0,8121	0,9318
Baseline ECOG-PS	0,5470	0,6186	0,4401	0,7246	0,9006	0,6060	0,7403	0,3893
Baseline Gewichtsgruppe	0,1226	0,4120	0,1199	0,1812	0,4967	0,6350	0,9111	0,6650
Baseline AFP	0,1783	0,1349	0,0849	**0,0414**	0,3810	0,2026	0,7226	0,1992
Ätiologie HBV	**0,0141**	0,2451	**0,0144**	0,5024	0,3345	**0,0294**	0,9994	**0,0008**
Ätiologie HCV	0,5150	0,3522	0,1887	0,9106	0,6196	0,9639	0,9299	**0,0169**
Ätiologie Alkohol	0,8465	0,9815	0,9287	0,6765	0,9551	0,3027	0,4567	0,2070
BCLC Staging	0,0996	0,1465	0,3220	0,9835	0,2059	0,4279	0,7903	0,3408
Leberzirrhose zu Baseline (CRF)	0,2897	0,9830	0,2529	0,8931	0,9881	0,2143	0,6713	0,5025
Leberzirrhose zu Baseline (IIR)	0,3596	0,7410	0,2889	0,7421	**0,0480**	0,1254	0,6400	0,9647
Vorbehandlung mit TACE	0,3151	0,4906	0,3155	0,3252	0,4722	0,3138	0,8413	0,0829

Subgruppen-Kategorie	Interaktions-p-Wert[a]							
	Fatigue	Übelkeit und Erbrechen	Schmerz	Dys-pnoe	Schlaf-losigkeit	Appetit-minderung	Obsti-pation	Diarrhö

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); TACE: Transarterielle Chemoembolisation

Quelle: SCO:SSiS 2018

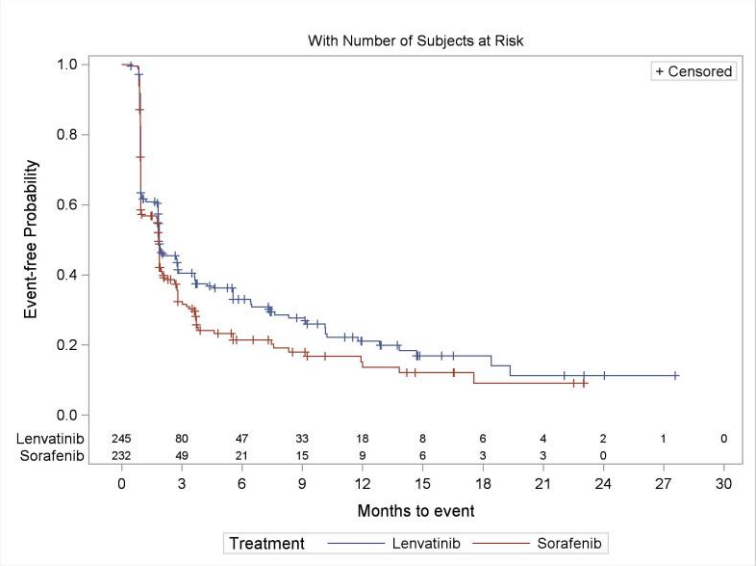


Abbildung 4-77: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

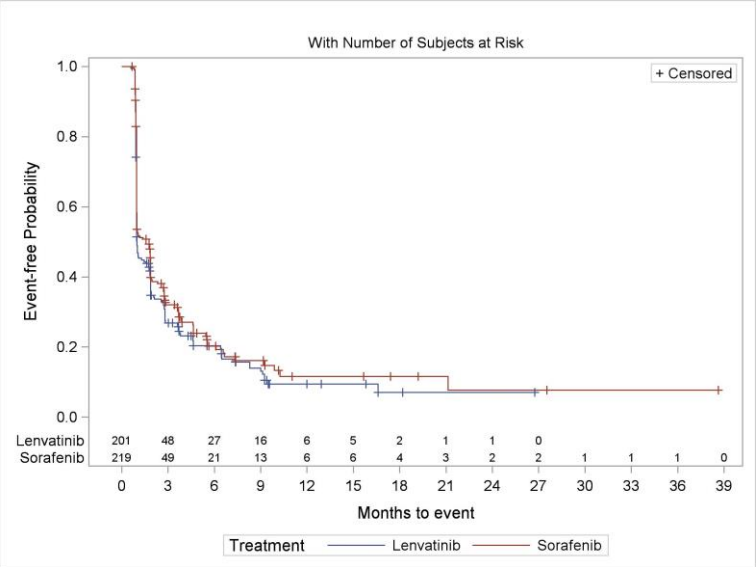


Abbildung 4-78: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fatigue** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

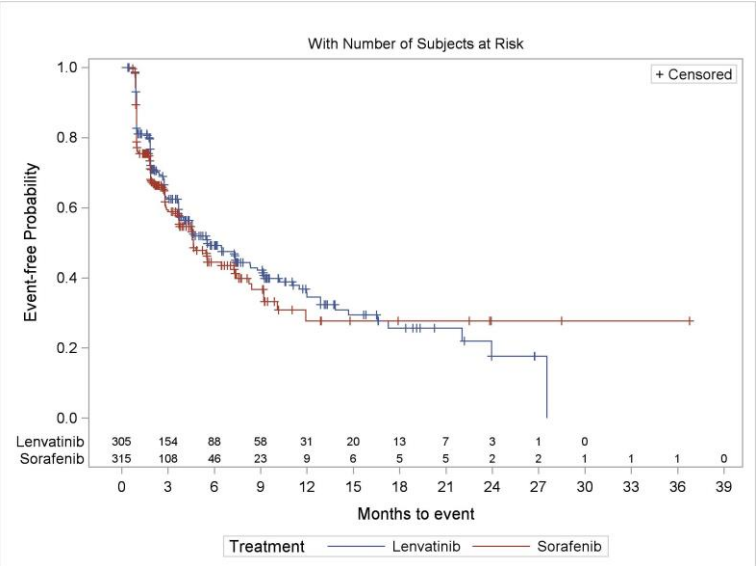


Abbildung 4-79: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Übelkeit** und **Erbrechen** des EORTC QLQ-C30 (MPVI oder ES oder beides „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

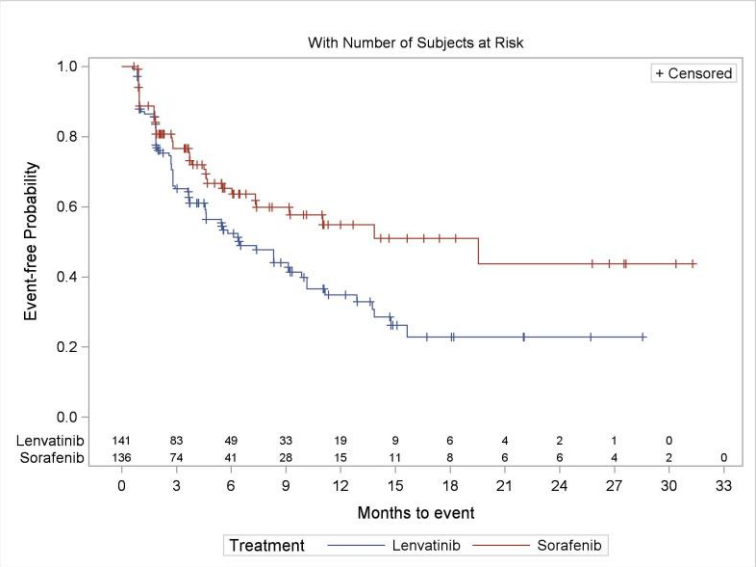


Abbildung 4-80: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Übelkeit** und **Erbrechen** des EORTC QLQ-C30 (MPVI oder ES oder beides „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

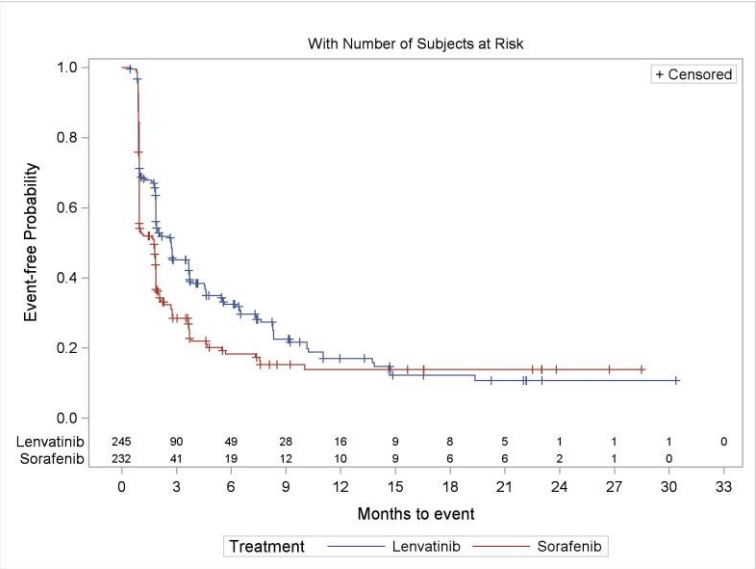


Abbildung 4-81: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

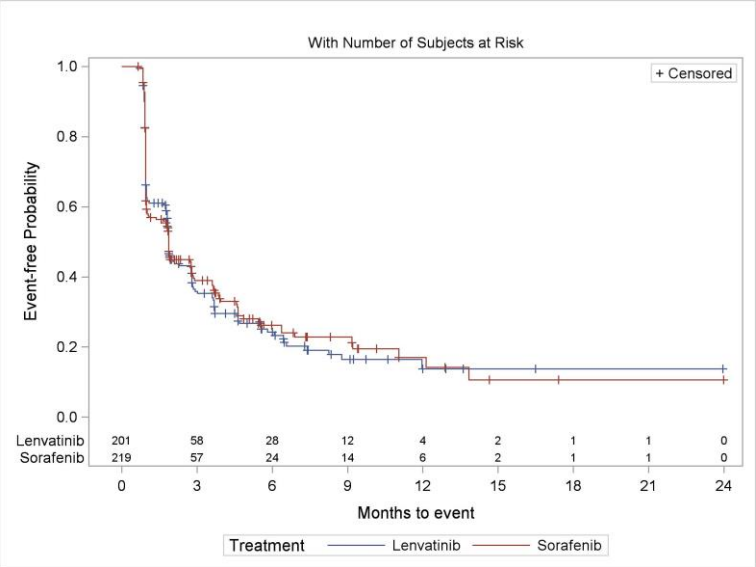


Abbildung 4-82: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

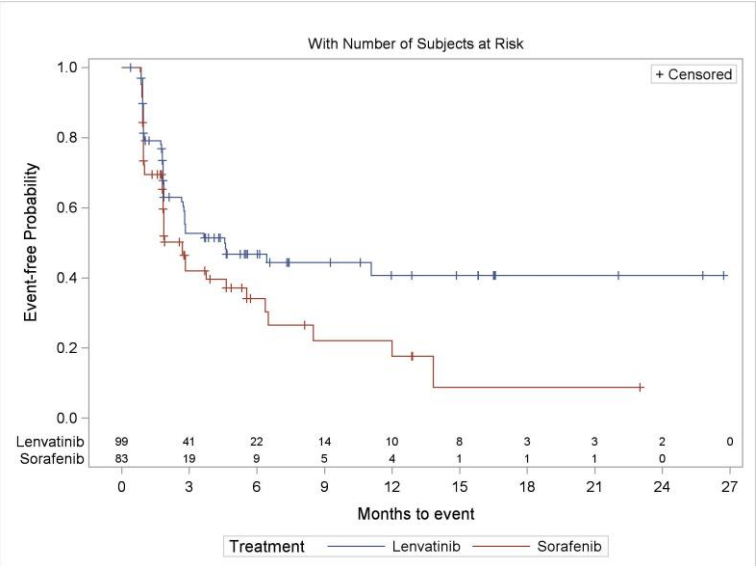


Abbildung 4-83: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (MPVI „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

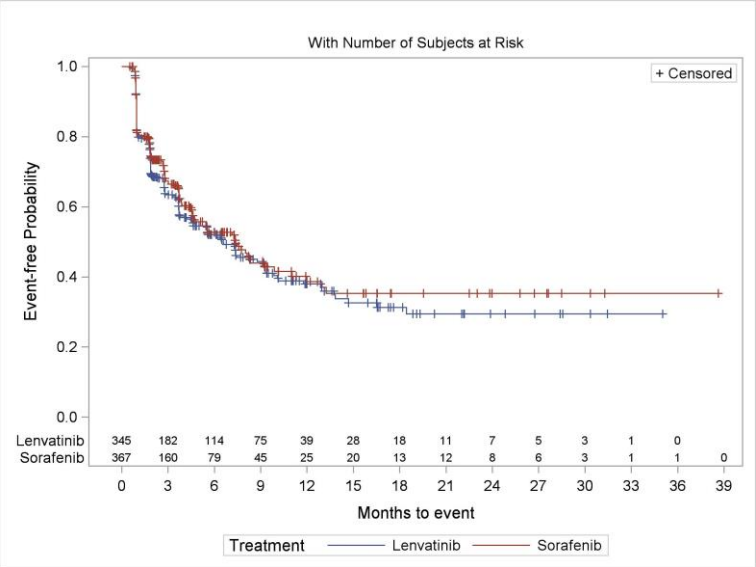


Abbildung 4-84: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (MPVI „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

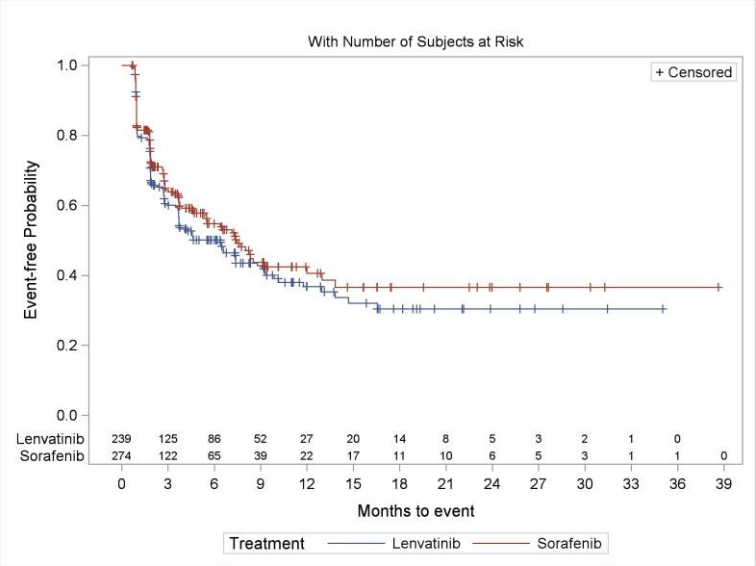


Abbildung 4-85: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (Baseline AFP „< 200 ng/mL“) aus Studie 304 mit dem zu bewertenden Arzneimittel

AFP: Alpha-Fetoprotein; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; mL: Milliliter; ng: Nanogramm
Quelle: SCO:SSiS 2018

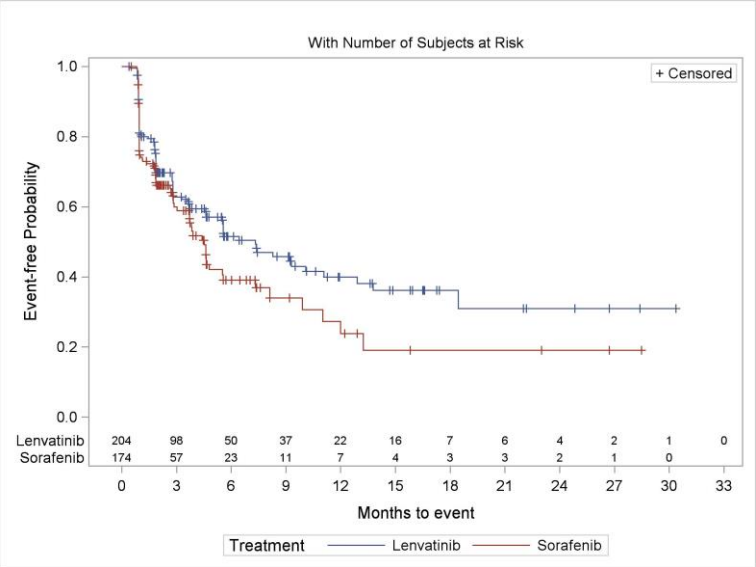


Abbildung 4-86: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Dyspnoe** des EORTC QLQ-C30 (Baseline AFP „≥ 200 ng/mL“) aus Studie 304 mit dem zu bewertenden Arzneimittel

AFP: Alpha-Fetoprotein; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; mL: Milliliter; ng: Nanogramm
Quelle: SCO:SSiS 2018

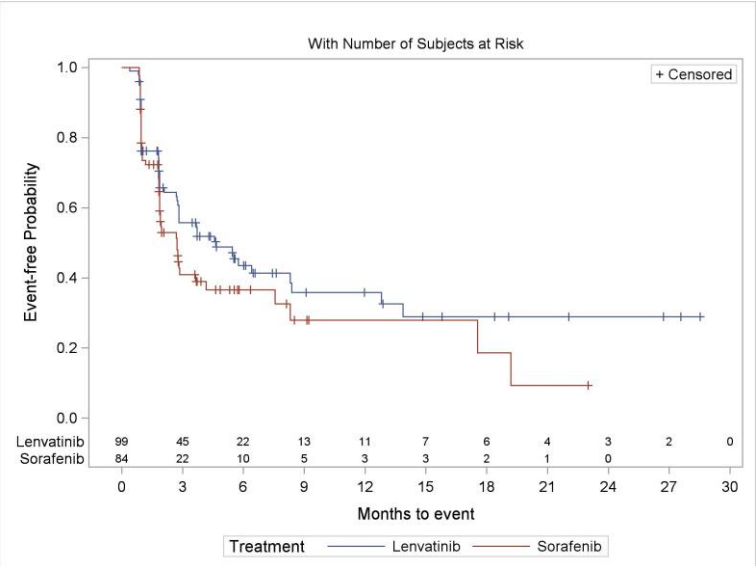


Abbildung 4-87: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (MPVI „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

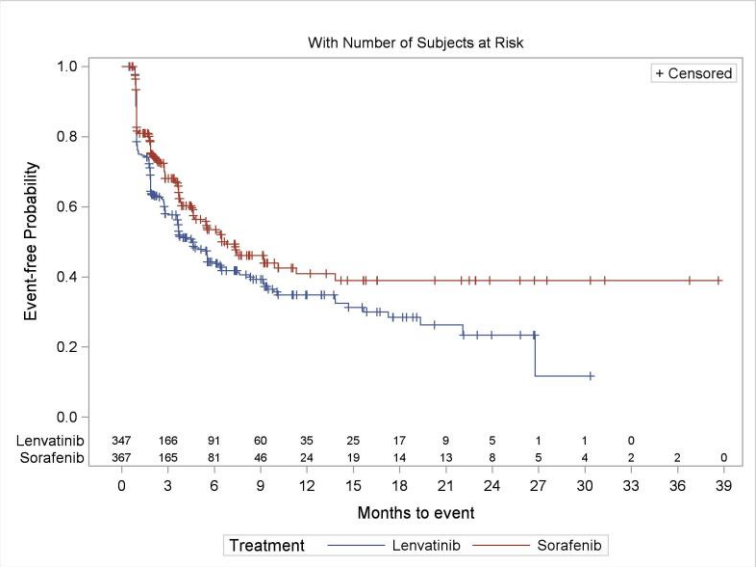


Abbildung 4-88: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (MPVI „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

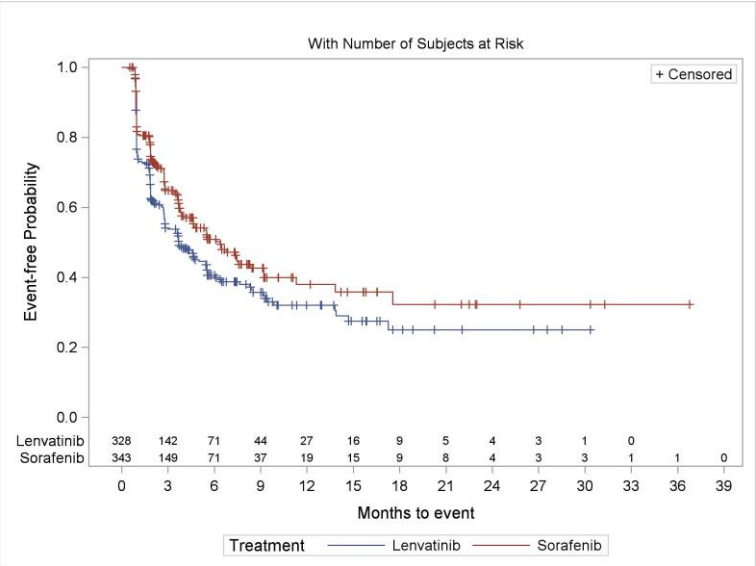


Abbildung 4-89: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IRR) „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; IIR: Independent Imaging Review (unabhängige radiologische Bewertung)
Quelle: SCO:SSiS 2018

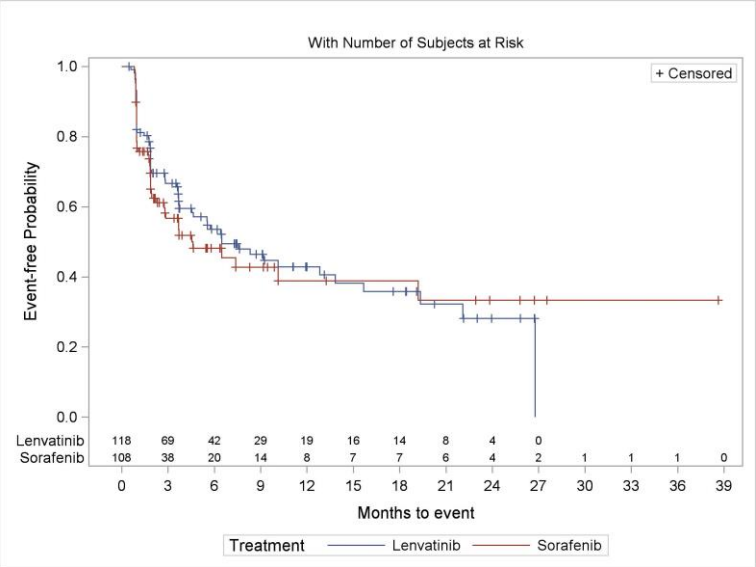


Abbildung 4-90: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schlaflosigkeit** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IRR) „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; IIR: Independent Imaging Review (unabhängige radiologische Bewertung)
Quelle: SCO:SSiS 2018

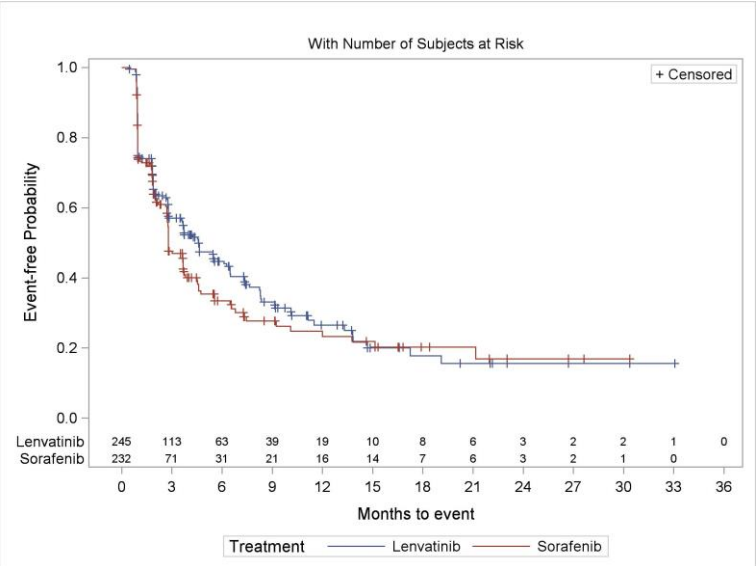


Abbildung 4-91: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

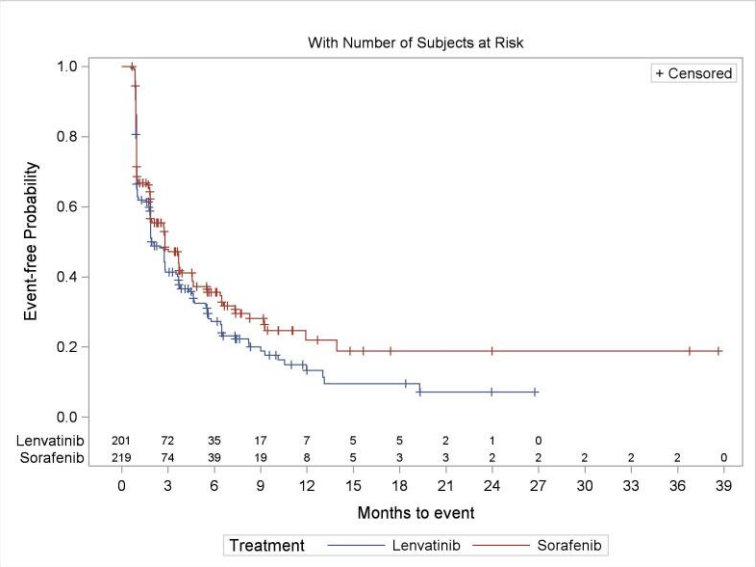
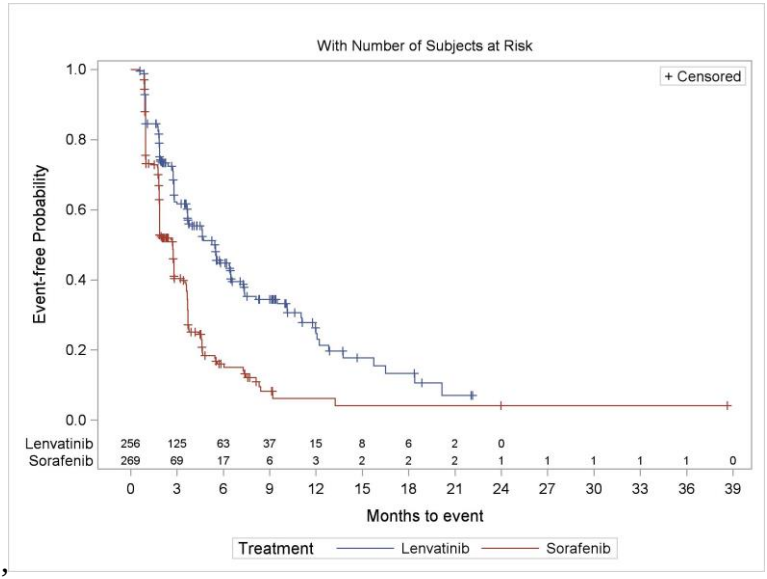


Abbildung 4-92: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Appetitminderung** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018



,

Abbildung 4-93: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Baseline Altersgruppe „< 65 Jahre) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

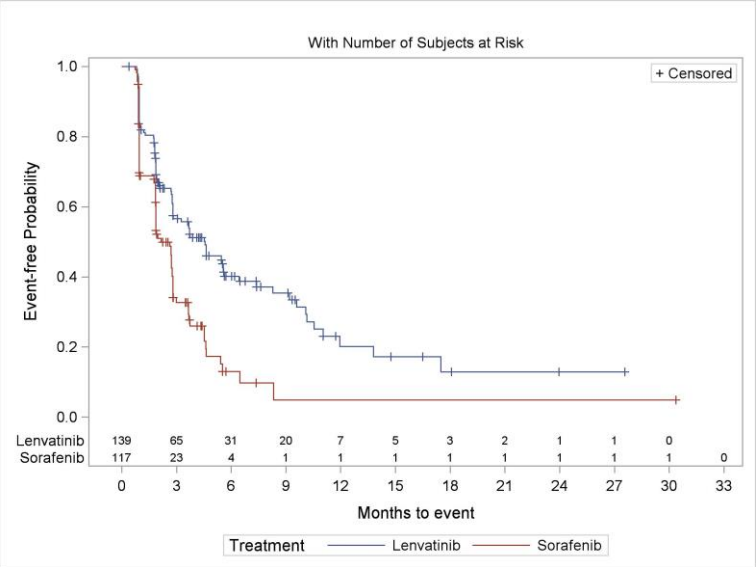


Abbildung 4-94: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Baseline Altersgruppe „≥ 65 bis < 75 Jahre") aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

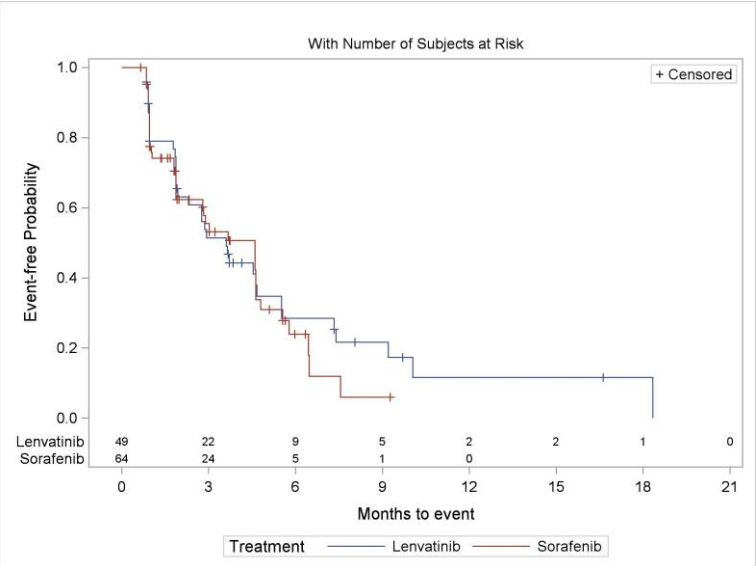


Abbildung 4-95: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Baseline Altersgruppe „≥ 75 Jahre") aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

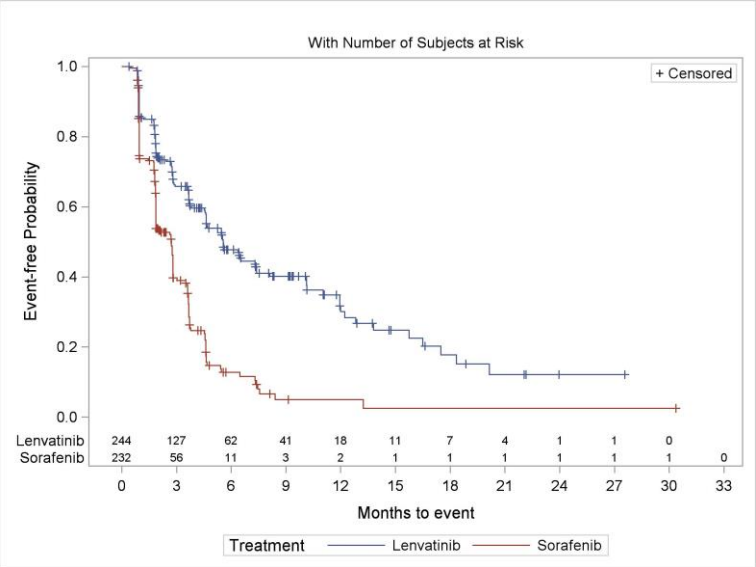


Abbildung 4-96: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HBV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

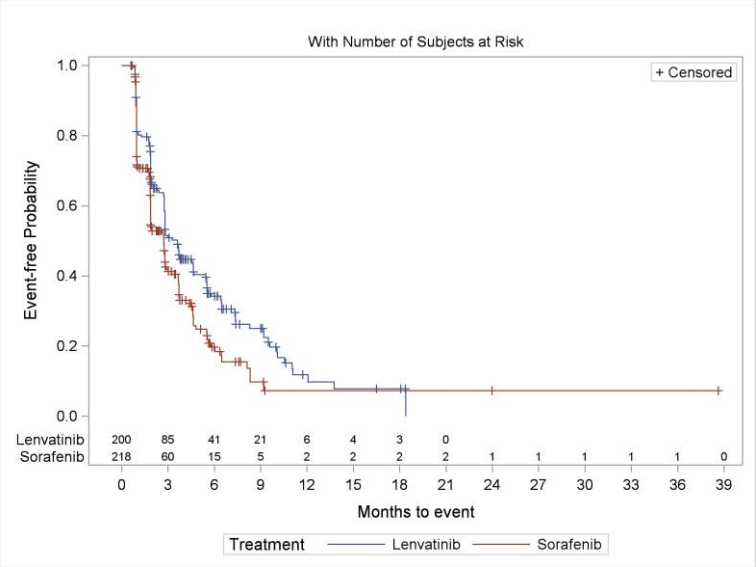


Abbildung 4-97: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HBV „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

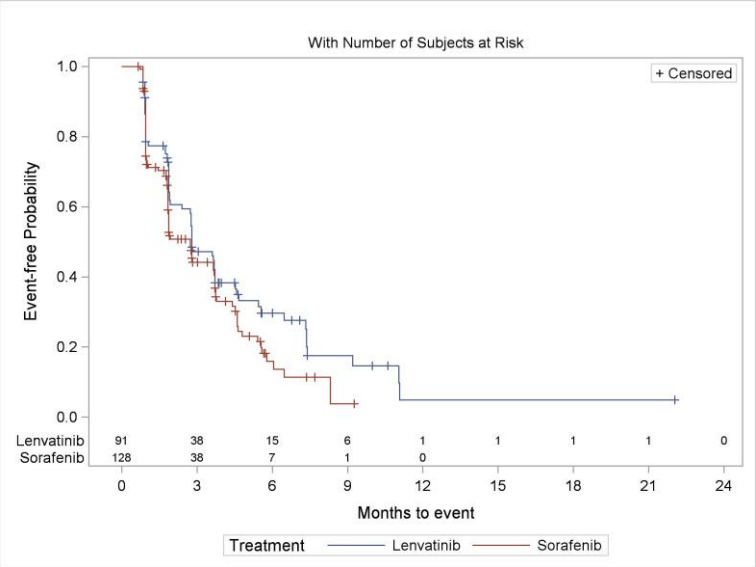


Abbildung 4-98: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HCV „Ja") aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HCV: Hepatitis C Virus
Quelle: SCO:SSiS 2018

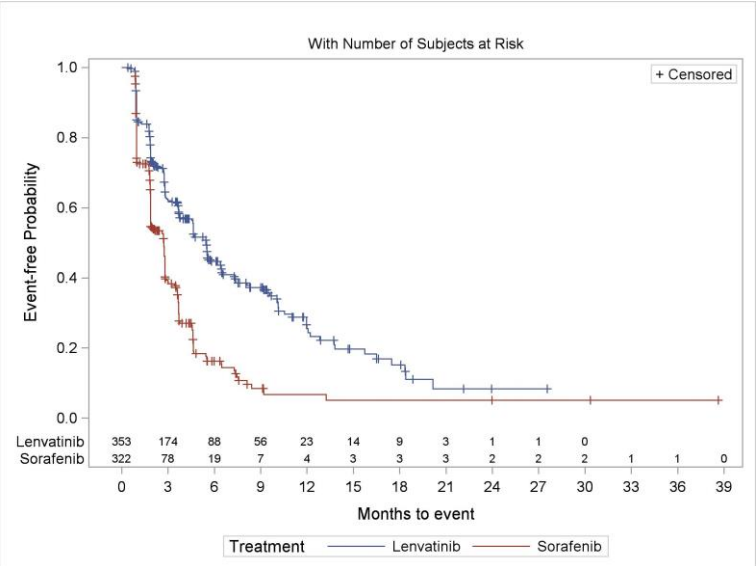


Abbildung 4-99: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Diarrhö** des EORTC QLQ-C30 (Ätiologie HCV „Nein") aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HCV: Hepatitis C Virus
Quelle: SCO:SSiS 2018

Tabelle 4-60: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (Symptome) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
EORTC QLQ-C30 Fatigue					
Alle	338/446 (75,8)	1,8 [1,05; 1,87]	330/451 (73,2)	1,8 [1,02; 1,87]	0,93 [0,797; 1,085] 0,3592
Ätiologie HBV					
Ja	171/245 (69,8)	1,9 [1,84; 2,76]	166/232 (71,6)	1,8 [1,74; 1,87]	0,80 [0,641; 0,985] 0,0356
Nein	167/201 (83,1)	1,0 [0,95; 1,77]	164/219 (74,9)	1,8 [0,99; 1,87]	1,16 [0,931; 1,434] 0,1892
EORTC QLQ-C30 Übelkeit und Erbrechen					
Alle	241/446 (54,0)	6,4 [4,60; 8,31]	189/451 (41,9)	7,3 [4,63; 9,20]	1,04 [0,857; 1,263] 0,6894
MPVI oder ES oder beides					
Ja	162/305 (53,1)	5,6 [4,53; 8,35]	144/315 (45,7)	4,6 [3,71; 7,29]	0,89 [0,712; 1,121] 0,3312
Nein	79/141 (56,0)	6,5 [4,60; 9,86]	45/136 (33,1)	19,5 [7,36; NE]	1,58 [1,094; 2,276] 0,0139
EORTC QLQ-C30 Schmerz					
Alle	321/446 (72,0)	1,9 [1,87; 2,76]	319/451 (70,7)	1,8 [1,12; 1,87]	0,81 [0,690; 0,949] 0,0091
Ätiologie HBV					
Ja	175/245 (71,4)	2,7 [1,87; 3,68]	170/232 (73,3)	1,8 [0,95; 1,87]	0,70 [0,565; 0,867] 0,0010
Nein	146/201 (72,6)	1,9 [1,81; 2,73]	149/219 (68,0)	1,9 [1,38; 2,76]	1,04 [0,825; 1,303] 0,7580
EORTC QLQ-C30 Dyspnoe					
Alle	230/444 (51,8)	6,4 [4,57; 8,28]	202/450 (44,9)	5,6 [4,57; 8,02]	0,97 [0,797; 1,170] 0,7217

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
MPVI					
Ja	48/99 (48,5)	4,6 [2,73; NE]	50/83 (60,2)	2,7 [1,84; 4,63]	0,65 [0,435; 0,967] 0,0323
Nein	182/345 (52,8)	6,6 [4,60; 9,20]	152/367 (41,4)	7,4 [4,80; 9,89]	1,09 [0,882; 1,359] 0,4095
Baseline AFP					
< 200 ng/mL	133/239 (55,6)	6,4 [3,68; 8,77]	118/274 (43,1)	7,5 [5,49; 11,99]	1,15 [0,900; 1,480] 0,2578
≥ 200 ng/mL	96/204 (47,1)	7,3 [4,63; 11,07]	84/174 (48,3)	4,5 [3,71; 5,55]	0,76 [0,567; 1,026] 0,0726
EORTC QLQ-C30 Schlaflosigkeit					
Alle	251/446 (56,3)	4,6 [3,68; 5,55]	199/451 (44,1)	6,1 [4,53; 7,56]	1,16 [0,958; 1,398] 0,1293
MPVI					
Ja	53/99 (53,5)	4,7 [2,76; 8,38]	50/84 (59,5)	2,7 [1,87; 4,17]	0,74 [0,504; 1,099] 0,1364
Nein	198/347 (57,1)	4,6 [3,65; 5,55]	149/367 (40,6)	6,8 [5,45; 10,12]	1,35 [1,091; 1,670] 0,0056
Leberzirrhose zu Baseline (IIR)					
Ja	188/328 (57,4)	3,7 [2,79; 5,45]	148/343 (43,1)	6,4 [4,63; 8,31]	1,33 [1,075; 1,655] 0,0085
Nein	63/118 (53,4)	6,5 [3,71; 13,83]	51/108 (47,2)	4,5 [2,76; 19,19]	0,89 [0,614; 1,293] 0,5434
EORTC QLQ-C30 Appetitminderung					
Alle	298/446 (66,8)	2,8 [2,76; 3,78]	268/451 (59,4)	2,8 [2,76; 3,68]	1,02 [0,858; 1,201] 0,8575
Ätiologie HBV					
Ja	148/245 (60,4)	4,6 [3,58; 6,47]	134/232 (57,8)	2,8 [2,73; 3,68]	0,85 [0,676; 1,082] 0,1919

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
Nein	150/201 (74,6)	1,9 [1,84; 2,79]	134/219 (61,2)	2,8 [1,87; 3,71]	1,24 [0,986; 1,571] 0,0661
EORTC QLQ-C30 Diarrhö					
Alle	266/444 (59,9)	4,6 [3,71; 5,52]	310/450 (68,9)	2,7 [1,87; 2,79]	0,52 [0,440; 0,623] <0,0001
Baseline Altersgruppe					
<65 Jahre	146/256 (57,0)	5,5 [3,75; 6,44]	190/269 (70,6)	2,7 [1,87; 2,79]	0,49 [0,392; 0,612] <0,0001
≥65 bis <75 Jahre	85/139 (61,2)	4,6 [2,79; 5,62]	81/117 (69,2)	2,2 [1,87; 2,76]	0,52 [0,376; 0,710] <0,0001
≥75 Jahre	35/49 (71,4)	3,6 [1,87; 4,67]	39/64 (60,9)	4,6 [1,87; 4,63]	0,88 [0,552; 1,419] 0,6114
Ätiologie HBV					
Ja	130/244 (53,3)	5,6 [4,63; 7,39]	163/232 (70,3)	2,7 [1,87; 2,79]	0,41 [0,320; 0,520] <0,0001
Nein	136/200 (68,0)	3,6 [2,79; 4,63]	147/218 (67,4)	2,7 [1,87; 2,83]	0,71 [0,561; 0,903] 0,0048
Ätiologie HCV					
Ja	65/91 (71,4)	2,8 [1,94; 3,71]	89/128 (69,5)	2,7 [1,84; 3,68]	0,76 [0,547; 1,050] 0,0943
Nein	201/353 (56,9)	5,5 [4,60; 6,41]	221/322 (68,6)	2,7 [1,87; 2,79]	0,49 [0,402; 0,597] <0,0001

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

c: Der p-Wert wurde anhand eines unstratifizierten Log-Rank-Tests berechnet.

AFP: Alpha-Fetoprotein; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); ITT: Intention-to-treat; KI: Konfidenzintervall; mL: Milliliter; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert; NE: not estimable (nicht berechenbar); ng: Nanogramm

Quelle: SCO:SSiS 2018

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Fatigue, gemessen auf der Skala des EORTC QLQ-C30, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Ätiologie HBV“ festgestellt (Tabelle 4-59). Für Patienten mit HBV als Ursache für ihre Erkrankung ist der Effekt gleichgerichtet wie in der Gesamtpopulation. Für alle anderen Patienten zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant (Tabelle 4-60). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Übelkeit und Erbrechen, gemessen auf der Skala des EORTC QLQ-C30, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „MPVI oder ES oder beides“ festgestellt (Tabelle 4-59). Für Patienten ohne eine vorliegende MPVI oder ES oder beides ist der Effekt gleichgerichtet wie in der Gesamtpopulation. Für Patienten mit einer vorliegenden MPVI oder ES oder beides zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant (Tabelle 4-60). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Schmerz, gemessen auf der Skala des EORTC QLQ-C30, wurde ebenfalls ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Ätiologie HBV“ festgestellt (Tabelle 4-59). Auch hier zeigt der Effekt für Patienten mit HBV als Ursache für ihre Erkrankung in die gleiche Richtung wie der Effekt in der Gesamtpopulation. Für alle anderen Patienten zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant (Tabelle 4-60). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Dyspnoe, gemessen auf der Skala des EORTC QLQ-C30, wurden Belege für Interaktion (Interaktions-p-Wert < 0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-59):

- MPVI
- Baseline AFP

Der Effekt für Patienten mit einer vorliegenden MPVI zeigt in die gleiche Richtung wie der Effekt in der Gesamtpopulation. Hingegen zeigt der Effekt für Patienten ohne eine vorliegende MPVI in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant. Bezüglich des AFP Werts zu Baseline ist der Effekt für Patienten mit einem Wert größer 200 ng/mL gleichgerichtet wie in der Gesamtpopulation. Der Effekt für Patienten mit einem Wert kleiner 200 ng/mL ist anders gerichtet, aber nicht statistisch signifikant (Tabelle 4-60). Somit konnte

kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Schlaflosigkeit, gemessen auf der Skala des EORTC QLQ-C30, Belege für Interaktion (Interaktions-p-Wert <0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-59):

- MPVI
- Leberzirrhose zu Baseline (IIR)

Für Patienten mit einer vorliegenden MPVI ist der Effekt anders gerichtet als in der Gesamtpopulation, aber nicht statistisch signifikant. Für Patienten ohne eine vorliegende MPVI zeigt der Effekt in die gleiche Richtung wie in der Gesamtpopulation (Tabelle 4-60). Für Patienten mit Leberzirrhose zu Baseline (IIR) war der Effekt gleichgerichtet wie in der Gesamtpopulation. Für Patienten ohne Leberzirrhose war der Effekt anders gerichtet, aber nicht statistisch signifikant (Tabelle 4-60). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Appetitminderung, gemessen auf der Skala des EORTC QLQ-C30, wurde ein Beleg für Interaktion (Interaktions-p-Wert <0,05) bezüglich des Subgruppenmerkmals „Ätiologie HBV“ festgestellt (Tabelle 4-59). Für Patienten mit HBV als Ursache für ihre Erkrankung zeigt der Effekt in die entgegengesetzte Richtung wie in der Gesamtpopulation, ist aber nicht statistisch signifikant. Für alle anderen Patienten ist der Effekt gleichgerichtet (Tabelle 4-60). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Diarrhö, gemessen auf der Skala des EORTC QLQ-C30, wurden Belege für Interaktion (Interaktions-p-Wert <0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-59):

- Baseline Altersgruppe
- Ätiologie HBV
- Ätiologie HCV

Bezüglich der aufgeführten Subgruppenmerkmale ist der Effekt in der jeweiligen Subgruppenausprägung gleichgerichtet wie in der Gesamtpopulation (Tabelle 4-60). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.1.3.2.2.3 Subgruppenanalyse für den Endpunkt EORTC QLQ-HCC18 (Symptome)

Tabelle 4-61: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	**Interaktions-p-Wert[a]**				
	Fatigue	**Ikterus**	**Schmerz**	**Fieber**	**Abdominale Schwellung**
Baseline Altersgruppe.	0,6316	0,4396	**0,0338**	0,3726	0,0950
Geschlecht	0,9963	0,9480	0,3877	0,8082	0,4930
Region	0,3912	0,6430	0,4159	0,1506	0,5489
MPVI	0,9261	0,2695	0,4910	0,1408	0,3137
ES	0,9225	0,7930	0,3150	0,5911	0,3535
MPVI oder ES oder beides	0,4854	0,4318	0,2574	0,2398	0,0512
Baseline ECOG-PS	0,9194	0,2846	0,5252	0,4079	0,1524
Baseline Gewichtsgruppe	0,2965	0,8562	0,8033	0,9995	0,9709
Baseline AFP	0,7118	0,1652	0,4124	**0,0244**	0,8786
Ätiologie HBV	0,0558	0,3221	0,1645	0,0632	0,6123
Ätiologie HCV	0,7940	0,7178	0,9743	0,6714	0,7112
Ätiologie Alkohol	0,8381	0,6413	0,6938	0,7564	0,8811
BCLC Staging	0,7168	0,2668	0,6745	0,5182	0,1380
Leberzirrhose zu Baseline (CRF)	0,8693	0,8521	0,6762	0,0673	0,2058
Leberzirrhose zu Baseline (IIR)	0,7345	0,6646	0,4620	0,2633	0,7775
Vorbehandlung mit TACE	0,3908	0,3727	0,1267	0,1919	0,5493

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); TACE: Transarterielle Chemoembolisation

Quelle: SCO:SSiS 2018

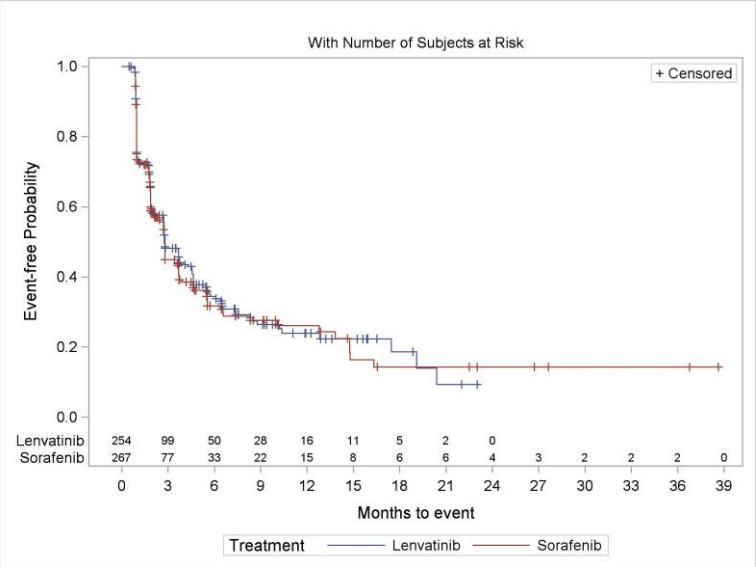


Abbildung 4-100: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 (Baseline Altersgruppe „< 65 Jahre“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

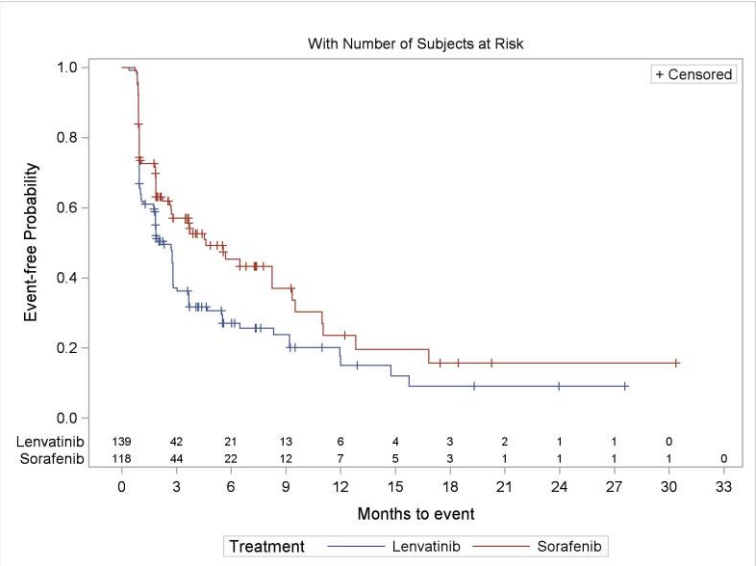


Abbildung 4-101: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 (Baseline Altersgruppe „≥ 65 bis < 75 Jahre“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

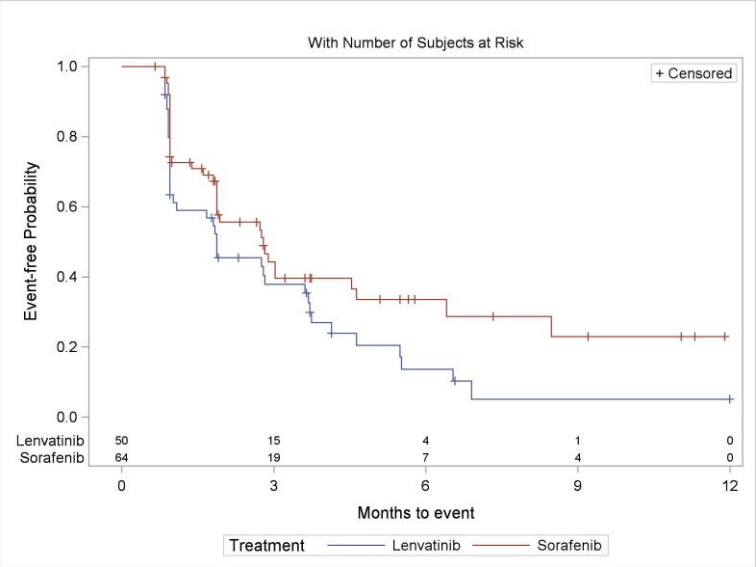


Abbildung 4-102: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Schmerz** des EORTC QLQ-HCC18 (Baseline Altersgruppe „≥ 75 Jahre“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

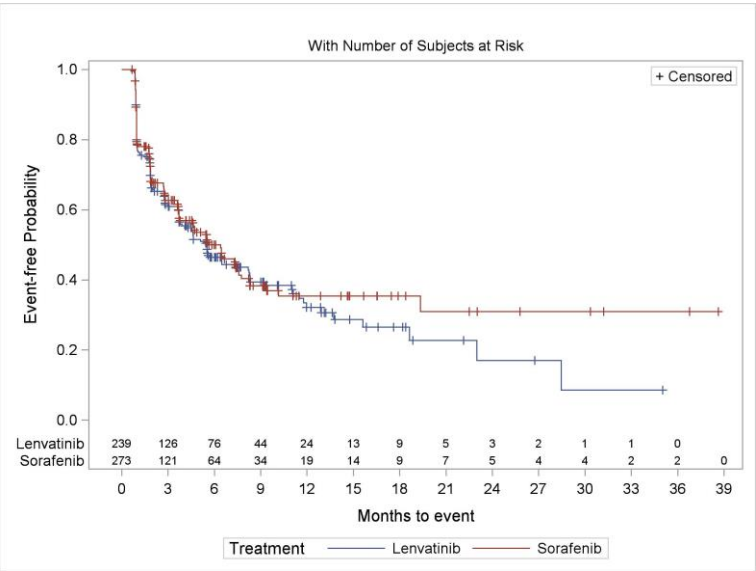


Abbildung 4-103: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 (Baseline AFP „< 200 ng/mL“) aus Studie 304 mit dem zu bewertenden Arzneimittel

AFP: Alpha-Fetoprotein; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; mL: Milliliter; ng: Nanogramm
Quelle: SCO:SSiS 2018

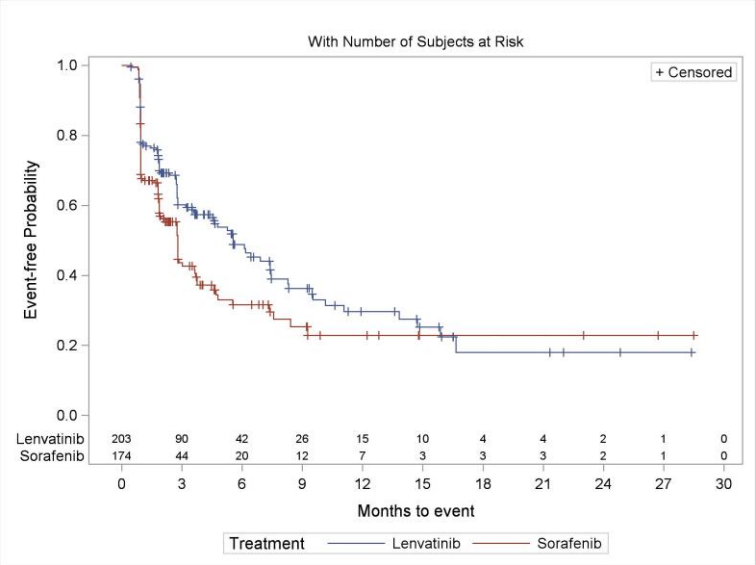


Abbildung 4-104: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Fieber** des EORTC QLQ-HCC18 (Baseline AFP „≥ 200 ng/mL“) aus Studie 304 mit dem zu bewertenden Arzneimittel

AFP: Alpha-Fetoprotein; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; mL: Milliliter; ng: Nanogramm
Quelle: SCO:SSiS 2018

Tabelle 4-62: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (Symptome) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
EORTC QLQ-HCC18 Schmerz					
Alle	303/443 (68,4)	2,8 [2,00; 2,79]	258/449 (57,5)	2,8 [2,73; 3,68]	1,14 [0,961; 1,349] 0,1315
Baseline Altersgruppe					
<65 Jahre	163/254 (64,2)	2,8 [2,73; 3,98]	158/267 (59,2)	2,8 [2,33; 3,65]	0,96 [0,768; 1,192] 0,6963
≥65 bis <75 Jahre	101/139 (72,7)	2,3 [1,84; 2,79]	63/118 (53,4)	4,6 [2,69; 8,25]	1,48 [1,078; 2,024] 0,0146

	Lenvatinib		Sorafenib		Lenvatinib vs. Sorafenib
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95%-KI]**	**n/N (%)**	**Median[a] [95%-KI]**	**HR[b] [95%-KI] p-Wert[c]**
≥75 Jahre	39/50 (78,0)	1,9 [0,95; 3,61]	37/64 (57,8)	2,8 [1,87; 4,63]	1,55 [0,984; 2,428] 0,0568
EORTC QLQ-HCC18 Fieber					
Alle	244/443 (55,1)	5,5 [4,57; 6,47]	231/449 (51,4)	3,7 [3,61; 5,52]	0,90 [0,752; 1,083] 0,2705
Baseline AFP					
< 200 ng/mL	138/239 (57,7)	5,5 [3,71; 8,18]	130/273 (47,6)	5,6 [4,60; 7,46]	1,09 [0,854; 1,381] 0,4989
≥ 200 ng/mL	105/203 (51,7)	5,6 [3,61; 7,43]	100/174 (57,5)	2,8 [1,87; 3,61]	0,70 [0,535; 0,929] 0,0127

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

c: Der p-Wert wurde anhand eines unstratifizierten Log-Rank-Tests berechnet.

AFP: Alpha-Fetoprotein; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; ITT: Intention-to-treat; KI: Konfidenzintervall; mL: Milliliter; n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert; ng: Nanogramm

Quelle: SCO:SSiS 2018

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Schmerz, gemessen auf der Skala des EORTC QLQ-HCC18, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Baseline Altersgruppe“ festgestellt (Tabelle 4-61). Für Patienten jünger als 65 Jahre ist der Effekt anders gerichtet als in der Gesamtpopulation, aber nicht statistisch signifikant. Für Patienten älter als 65 Jahre zeigt der Effekt in die gleiche Richtung (Tabelle 4-62). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Symptoms Fieber, gemessen auf der Skala des EORTC QLQ-HCC18, wurde ebenfalls ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Baseline AFP“ festgestellt (Tabelle 4-61). Hier zeigt der Effekt für Patienten mit einem Wert kleiner 200 ng/mL in die entgegengesetzte Richtung wie der Effekt in der Gesamtpopulation, ist aber nicht statistisch signifikant. Für Patienten mit einem Wert größer 200 ng/mL zeigt der Effekt in die gleiche Richtung (Tabelle 4-62). Somit konnte kein für die Nutzenbewertung relevanter

Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.1.3.2.2.4 Subgruppenanalyse für den Endpunkt EQ-5D VAS

Tabelle 4-63: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	**Interaktions-p-Wert[a]**	
	EQ-5D VAS, MID=7mm	**EQ-5D VAS, MID=10mm**
Baseline Altersgruppe,	0,5934	0,4213
Geschlecht	0,3410	0,3048
Region	0,4761	0,5949
MPVI	0,0729	0,1044
ES	0,4871	0,4278
MPVI oder ES oder beides	0,0701	**0,0399**
Baseline ECOG-PS	0,1480	0,1255
Baseline Gewichtsgruppe	0,2718	0,3827
Baseline AFP	0,0596	0,1554
Ätiologie HBV	0,5649	0,7142
Ätiologie HCV	0,6614	0,8899
Ätiologie Alkohol	0,9039	0,9904
BCLC Staging	0,5925	0,5841
Leberzirrhose zu Baseline (CRF)	0,2925	0,3695
Leberzirrhose zu Baseline (IIR)	0,7801	0,5009
Vorbehandlung mit TACE	0,8748	0,9115

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MID: Minimal Important Difference; mm: Millimeter; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); TACE: Transarterielle Chemoembolisation

Quelle: SCO:SSiS 2018

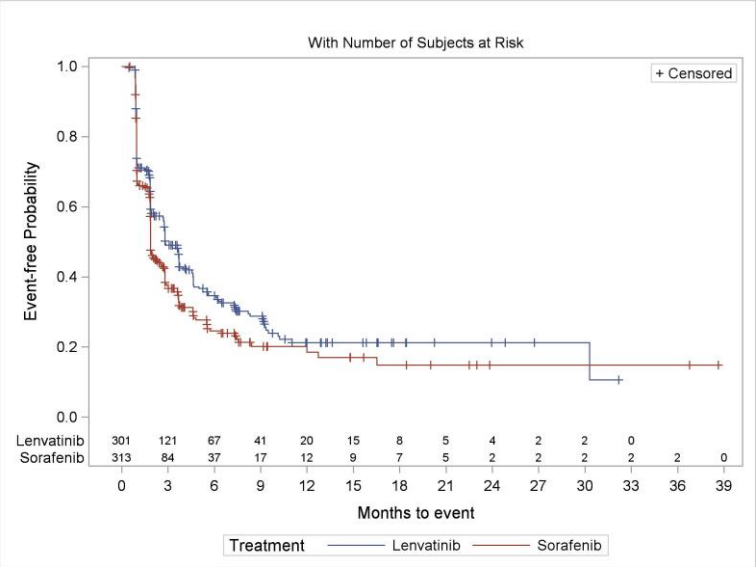


Abbildung 4-105: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS (MPVI oder ES oder beides „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; ES: Extrahepatische Ausbreitung (Extrahepatic Spread); MID: Minimal Important Difference; mm: Millimeter; MPVI: Makroskopische Invasion der Pfortader (Macroscopic Portal Vein Invasion)
Quelle: SCO:SSiS 2018

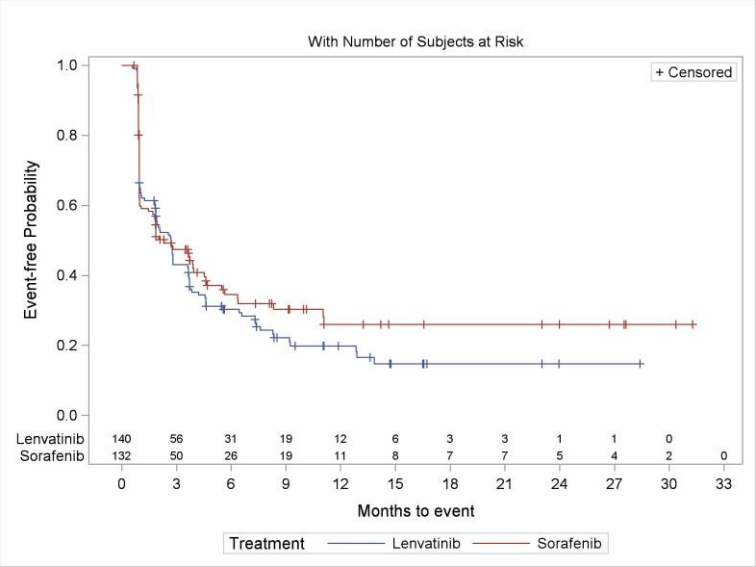


Abbildung 4-106: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS (MPVI oder ES oder beides „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; ES: Extrahepatische Ausbreitung (Extrahepatic Spread); MID: Minimal Important Difference; mm: Millimeter; MPVI: Makroskopische Invasion der Pfortader (Macroscopic Portal Vein Invasion)
Quelle: SCO:SSiS 2018

Tabelle 4-64: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der EQ-5D VAS stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95% KI]**	**n/N (%)**	**Median[a] [95% KI]**	**HR[b] [95% KI] p-Wert[c]**
EQ-5D VAS, MID=10mm					
Alle	303/441 (68,7)	2,8 [2,69; 3,68]	295/445 (66,3)	1,9 [1,87; 2,56]	0,89 [0,758; 1,050] 0,1713
MPVI oder ES oder beides					
Ja	196/301 (65,1)	2,8 [2,73; 3,71]	212/313 (67,7)	1,9 [1,87; 2,37]	0,79 [0,646; 0,955] 0,0154
Nein	107/140 (76,4)	2,7 [1,87; 3,65]	83/132 (62,9)	2,3 [1,45; 3,94]	1,14 [0,858; 1,523] 0,3595

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt

b: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet

c: Der p-Wert wurde anhand eines unstratifizierten Log-Rank-Tests berechnet.

EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; ES: Extrahepatische Ausbreitung (Extrahepatic Spread); HR: Hazard Ratio; KI: Konfidenzintervall; MID: Minimal Important Difference; mm: Millimeter; MPVI: Makroskopische Invasion der Pfortader (Macroscopic Portal Vein Invasion); n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline- und mind. einem Folgewert

Quelle: SCO:SSiS 2018

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 mm auf der EQ-5D VAS ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „MPVI oder ES oder beides“ festgestellt (Tabelle 4-63). Für Patienten mit einer vorliegenden MPVI oder ES oder beides ist der Effekt gleichgerichtet wie in der Gesamtpopulation. Für Patienten ohne MPVI oder ES oder beides zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant (Tabelle 4-64). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.1.3.2.3 Subgruppenanalyse für den Endpunkt Gesundheitsbezogene Lebensqualität

4.3.1.3.2.3.1 Subgruppenanalyse für die EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität)

Tabelle 4-65: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	**Interaktions-p-Wert[a]**						
	Physische Funktion	**Rollen-funktions-niveau**	**Emotionale Funktion**	**Kognitive Funktion**	**Soziale Funktion**	**Finanzielle Schwierig-keiten**	**Globaler Gesund-heits-status**
Baseline Altersgruppe	0,6847	0,7239	0,9591	0,5627	0,7359	0,2430	0,2780
Geschlecht	0,2456	0,4102	0,5973	0,5035	0,5758	0,1631	0,5283
Region	0,5367	0,2435	**0,0008**	**0,0123**	0,6729	0,4623	0,2364
MPVI	0,0915	0,4847	**0,0323**	0,0576	0,7258	0,9220	0,9343
ES	0,9512	0,6081	0,2656	0,7064	0,3957	0,6744	0,4261
MPVI oder ES oder beides	0,0918	0,6327	**0,0013**	0,0900	0,2934	0,9637	0,1939
Baseline ECOG-PS	0,7152	0,0900	0,1367	0,0755	0,6442	0,2546	0,2595
Baseline Gewichtsgruppe	0,3824	0,1011	0,4528	0,1059	0,9154	0,5336	**0,0452**
Baseline AFP	0,4787	0,4161	0,2481	0,1813	0,3738	0,5512	0,7199
Ätiologie HBV	**0,0313**	0,0591	**0,0360**	0,5433	0,1012	0,0830	**0,0232**
Ätiologie HCV	0,6441	0,9969	0,8205	0,7512	0,5861	0,6649	0,8286
Ätiologie Alkohol	0,6360	0,9835	0,4380	0,1868	0,7295	0,1150	0,4304
BCLC Staging	0,1729	0,6166	**0,0393**	0,3889	0,5983	0,7652	0,5164
Leberzirrhose zu Baseline (CRF)	0,6848	0,9449	0,4244	0,7723	0,7248	0,9514	0,6148
Leberzirrhose zu Baseline (IIR)	0,3456	0,1746	0,3404	0,7062	**0,0118**	0,9097	0,9686

Subgruppen-Kategorie	Interaktions-p-Wert[a]						
	Physische Funktion	Rollen-funktions-niveau	Emotionale Funktion	Kognitive Funktion	Soziale Funktion	Finanzielle Schwierig-keiten	Globaler Gesund-heits-status
Vorbehandlung mit TACE	0,3902	0,2167	0,6630	0,0524	0,8042	0,8393	0,1566

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); TACE: Transarterielle Chemoembolisation

Quelle: SCO:SSiS 2018

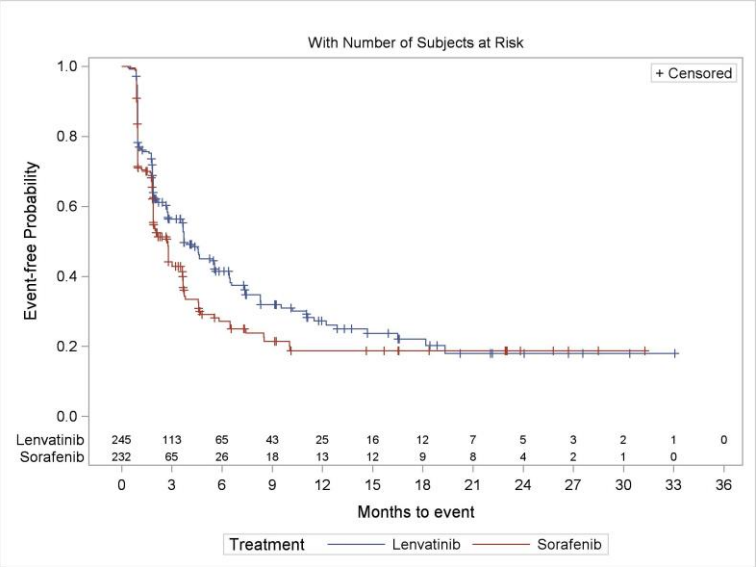


Abbildung 4-107: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus

Quelle: SCO:SSiS 2018

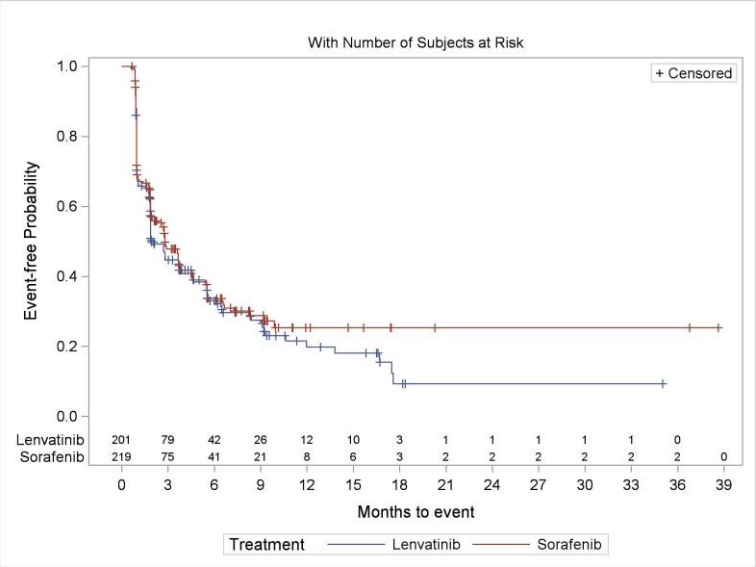


Abbildung 4-108: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Physische Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

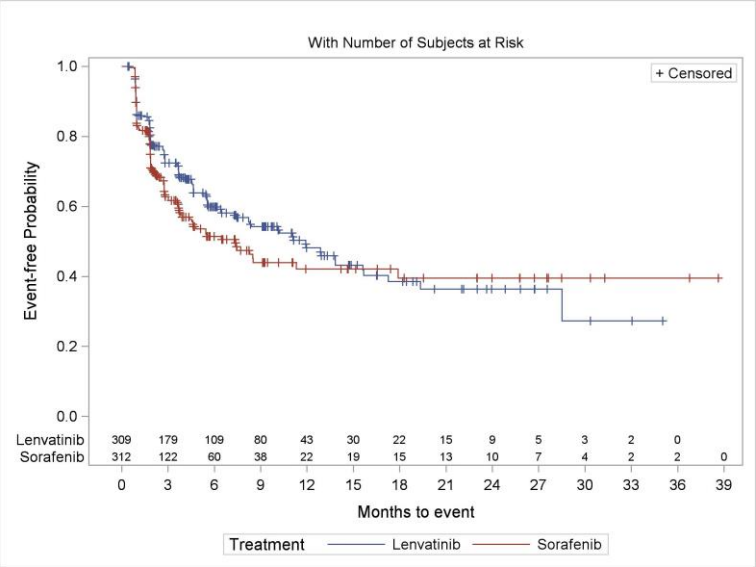


Abbildung 4-109: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Region[a] „Asien-Pazifik“) aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

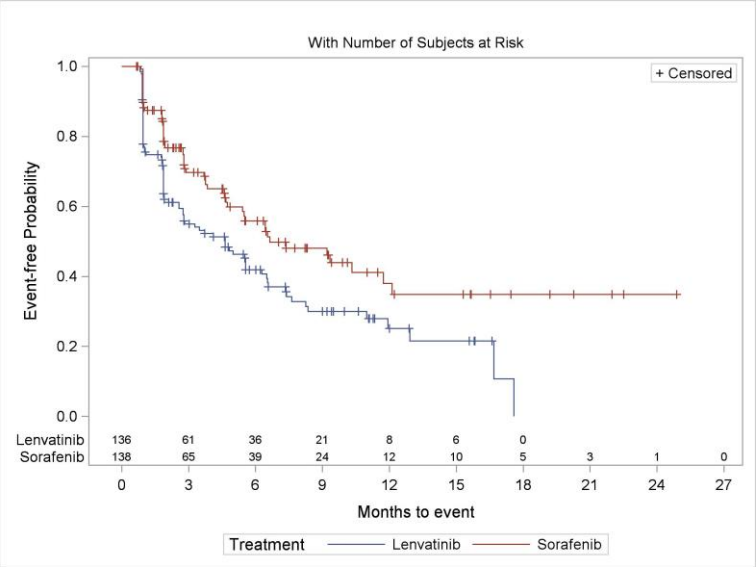


Abbildung 4-110: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Region[a] „westliche Regionen“) aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

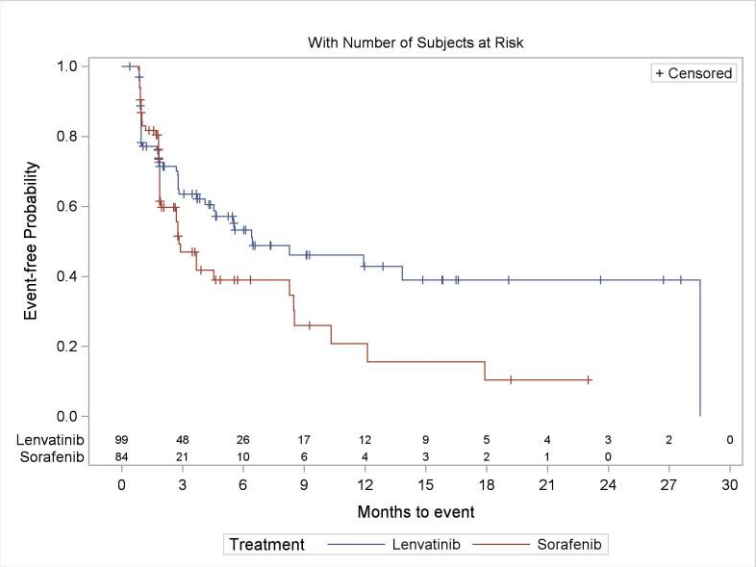


Abbildung 4-111: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

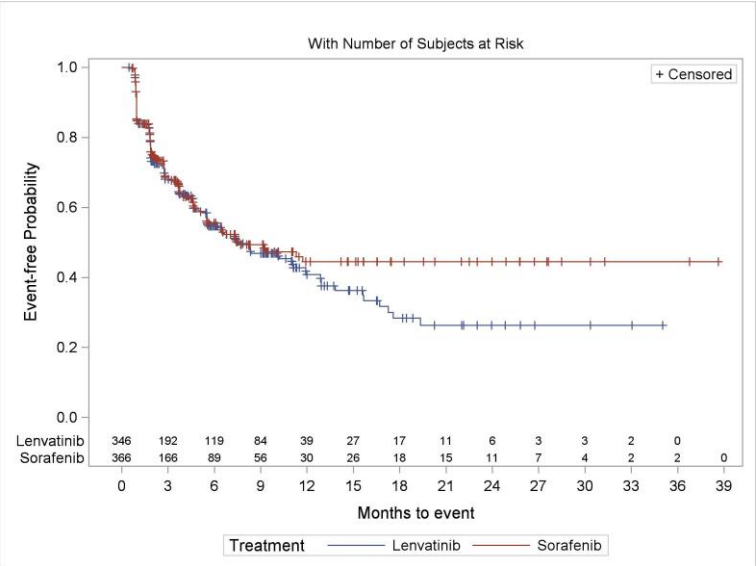


Abbildung 4-112: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

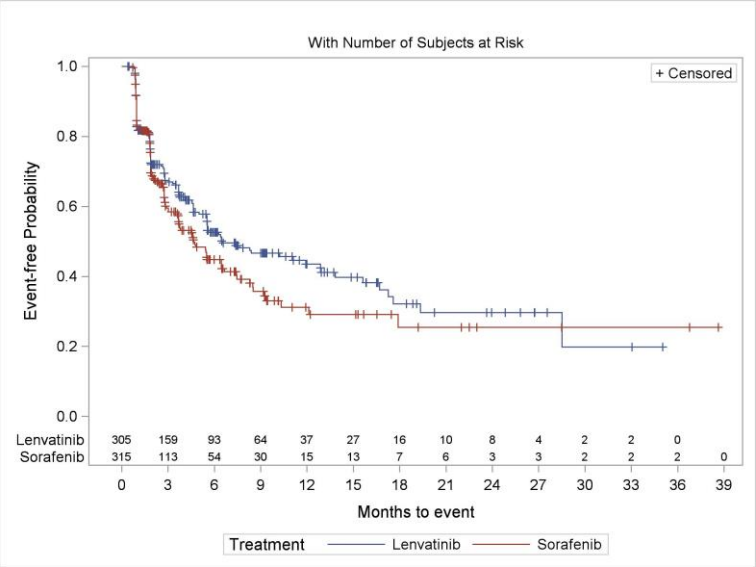


Abbildung 4-113: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI oder ES oder beides „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatische Ausbreitung (Extrahepatic Spread); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

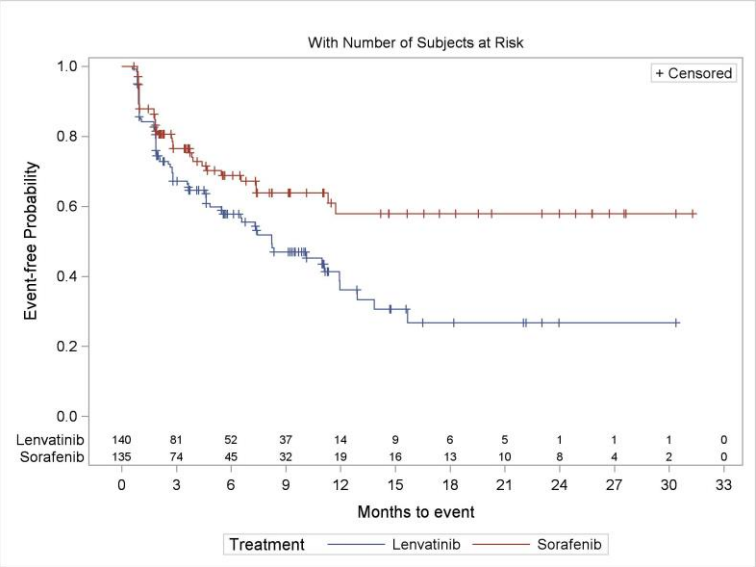


Abbildung 4-114: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (MPVI oder ES oder beides „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatische Ausbreitung (Extrahepatic Spread); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

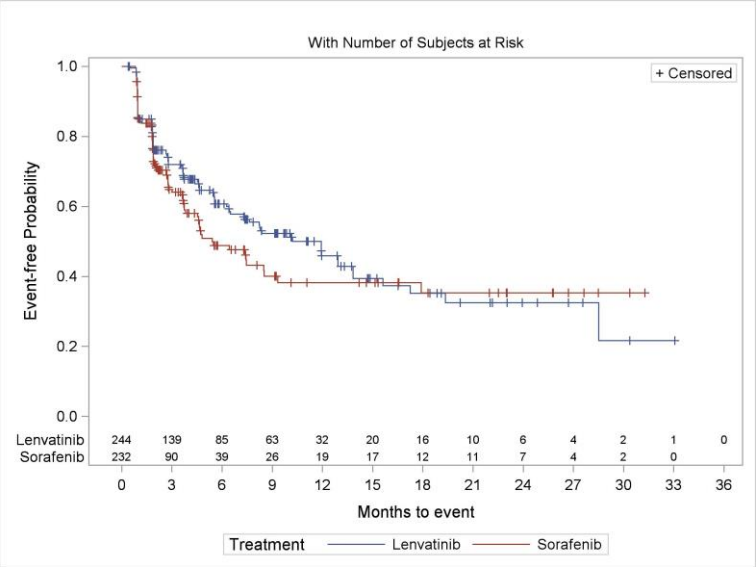


Abbildung 4-115: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

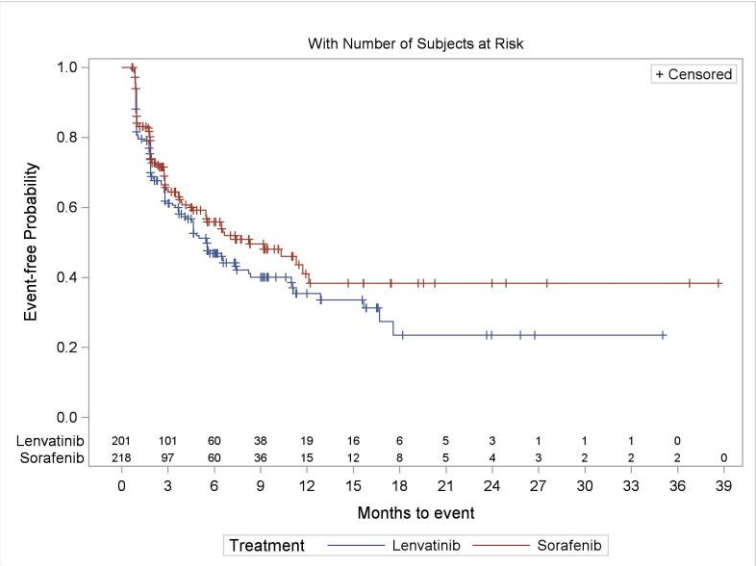


Abbildung 4-116: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

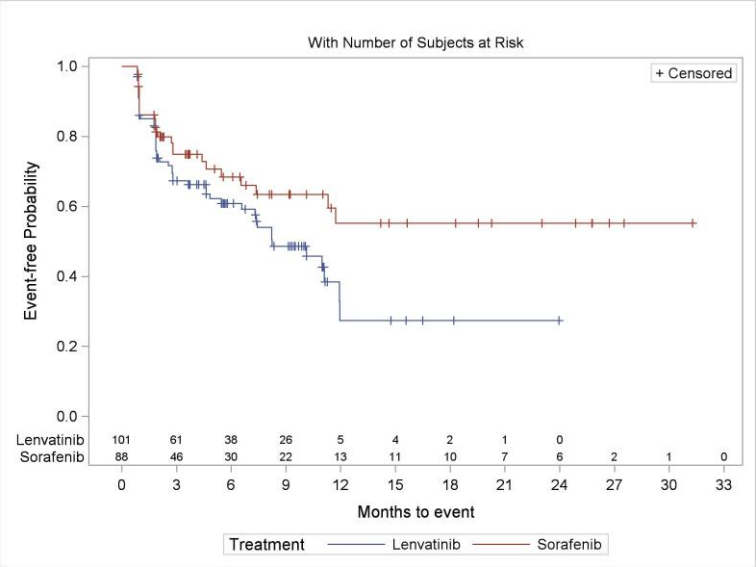


Abbildung 4-117: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (BCLC Staging „Stadium B“) aus Studie 304 mit dem zu bewertenden Arzneimittel

BCLC: Barcelona Clinic Liver Cancer; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

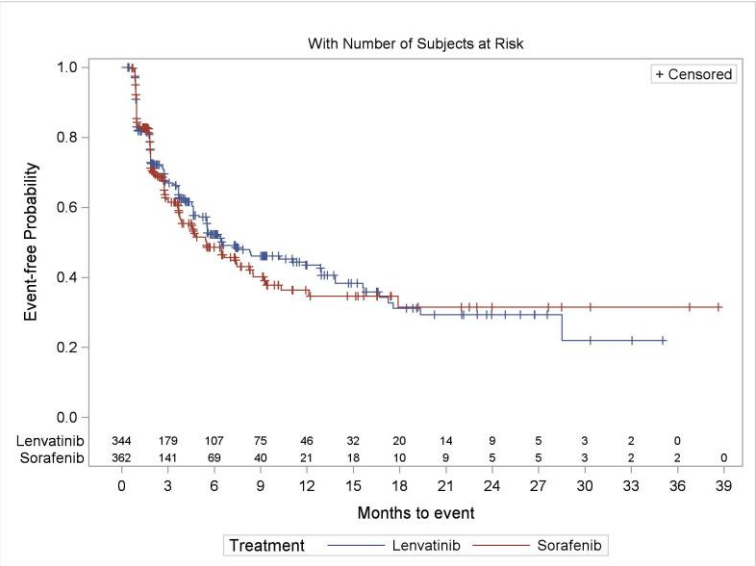


Abbildung 4-118: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Emotionale Funktion** des EORTC QLQ-C30 (BCLC Staging „Stadium C“) aus Studie 304 mit dem zu bewertenden Arzneimittel

BCLC: Barcelona Clinic Liver Cancer; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

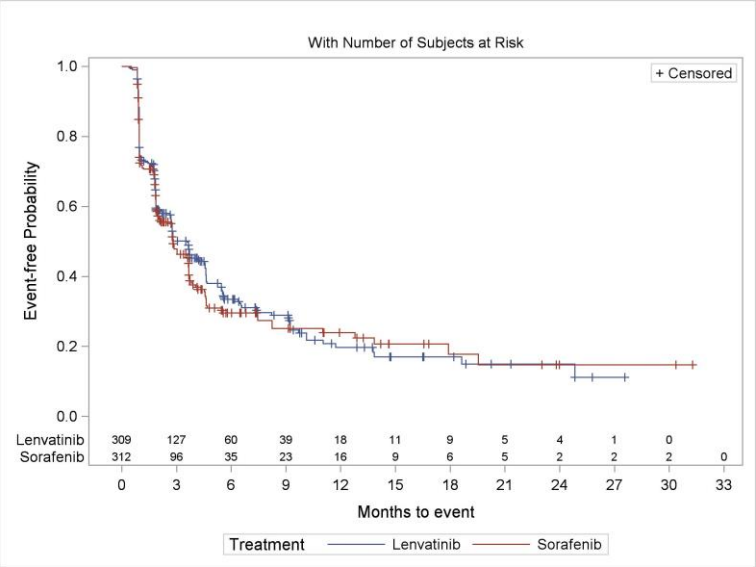


Abbildung 4-119: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 (Region[a] „Asien-Pazifik“) aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

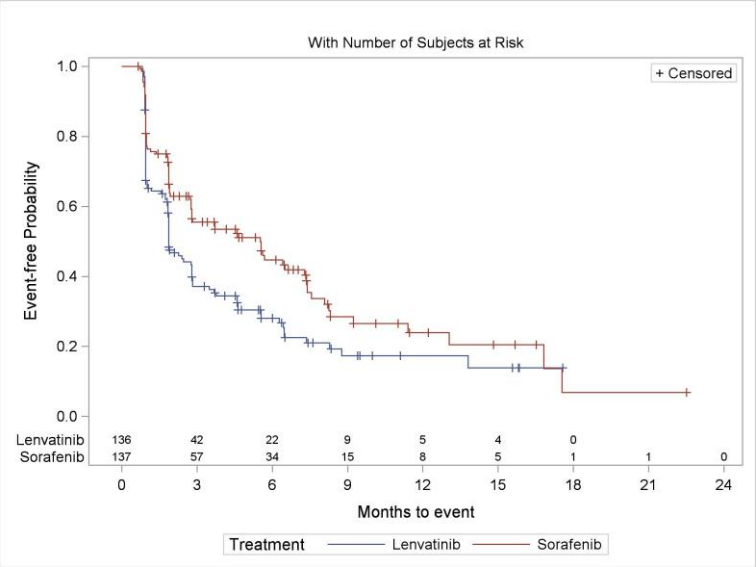


Abbildung 4-120: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Kognitive Funktion** des EORTC QLQ-C30 (Region[a] „Westliche Regionen“) aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

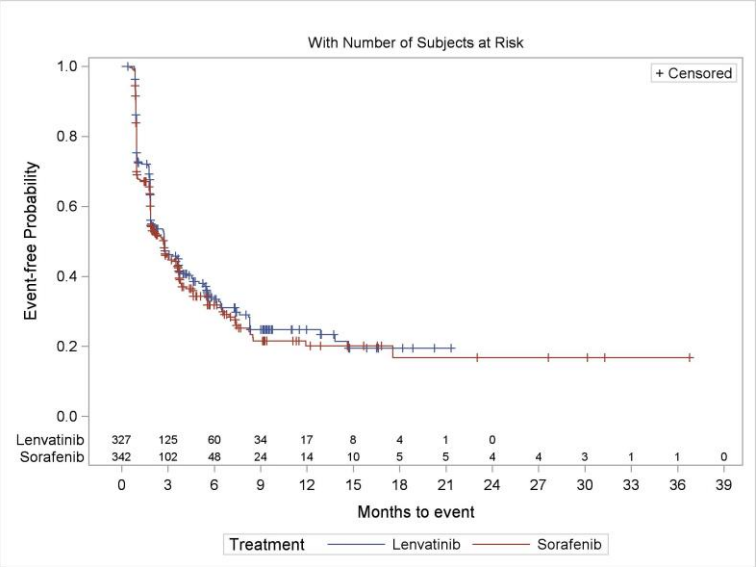


Abbildung 4-121: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IIR) „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; IIR: Independent Imaging Review (unabhängige radiologische Bewertung)
Quelle: SCO:SSiS 2018

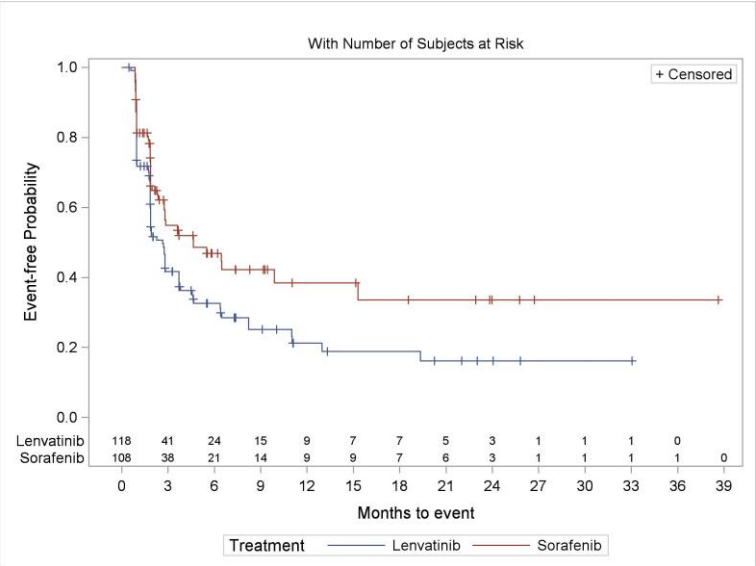


Abbildung 4-122: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Soziale Funktion** des EORTC QLQ-C30 (Leberzirrhose zu Baseline (IIR) „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; IIR: Independent Imaging Review (unabhängige radiologische Bewertung)
Quelle: SCO:SSiS 2018

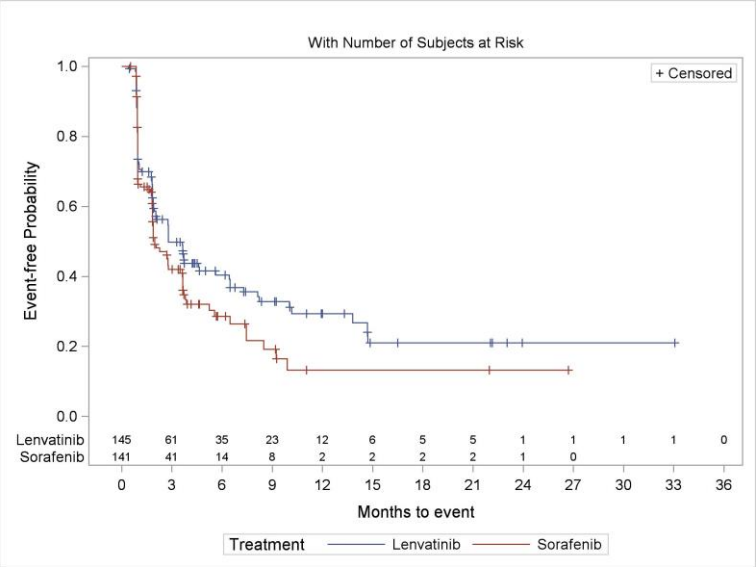


Abbildung 4-123: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Baseline Gewichtsgruppe „< 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; kg: Kilogramm
Quelle: SCO:SSiS 2018

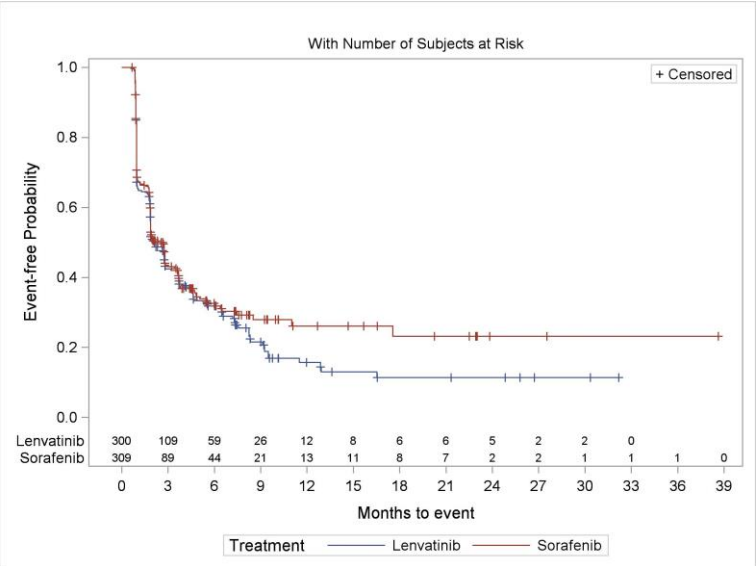


Abbildung 4-124: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Baseline Gewichtsgruppe „≥ 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; kg: Kilogramm
Quelle: SCO:SSiS 2018

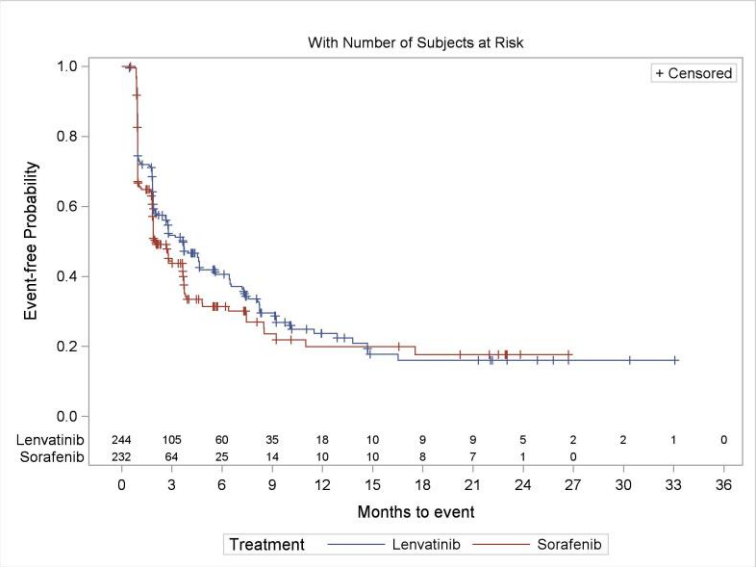


Abbildung 4-125: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

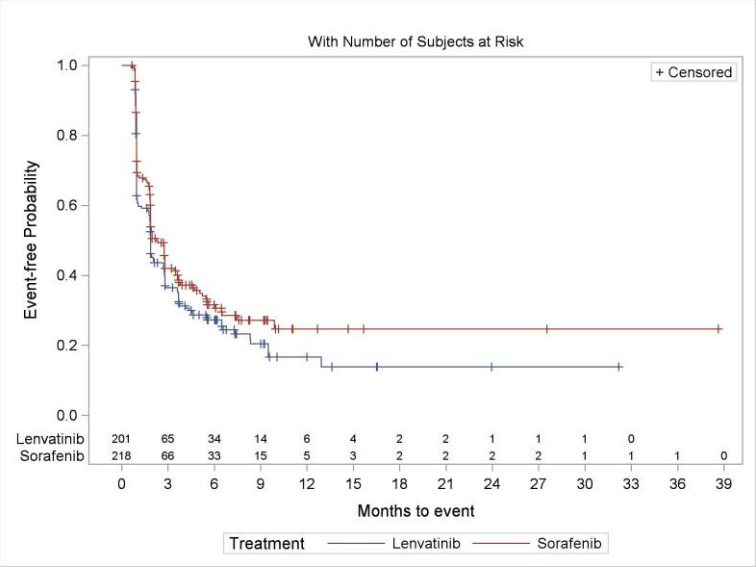


Abbildung 4-126: Subgruppenanalyse: Kaplan-Meier-Kurve der Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Globaler Gesundheitsstatus** des EORTC QLQ-C30 (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; HBV: Hepatitis B Virus
Quelle: SCO:SSiS 2018

Tabelle 4-66: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-C30 (gesundheitsbezogene Lebensqualität) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[b] [95% KI]**	**n/N (%)**	**Median[b] [95% KI]**	**HR[c] [95% KI] p-Wert[d]**
EORTC QLQ-C30 Physische Funktion					
Alle	294/446 (65,9)	3,7 [2,69; 4,53]	277/451 (61,4)	2,8 [1,97; 3,61]	0,89 [0,756; 1,056] 0,1870
Ätiologie HBV					
Ja	154/245 (62,9)	3,7 [2,79; 5,55]	145/232 (62,5)	2,8 [1,87; 3,02]	0,76 [0,605; 0,956] 0,0189
Nein	140/201 (69,7)	1,9 [1,87; 3,71]	132/219 (60,3)	2,8 [1,97; 3,84]	1,11 [0,877; 1,413] 0,3770
EORTC QLQ-C30 Emotionale Funktion					
Alle	216/445 (48,5)	7,4 [5,55; 11,07]	187/450 (41,6)	6,6 [4,80; 9,33]	1,01 [0,827; 1,235] 0,9142
Region[a]					
Asien-Pazifik	131/309 (42,4)	11,9 [8,21; 15,61]	129/312 (41,3)	7,4 [4,40; 17,91]	0,81 [0,634; 1,035] 0,0918
Westliche Regionen	85/136 (62,5)	4,6 [2,76; 6,28]	58/138 (42,0)	6,6 [4,73; 12,12]	1,64 [1,173; 2,293] 0,0034
MPVI					
Ja	45/99 (45,5)	6,4 [4,11; 28,52]	45/84 (53,6)	2,8 [1,87; 8,28]	0,66 [0,431; 1,003] 0,0500
Nein	171/346 (49,4)	7,6 [5,55; 11,10]	142/366 (38,8)	7,5 [5,49; NE]	1,11 [0,887; 1,386] 0,3617
MPVI oder ES oder beides					
Ja	145/305 (47,5)	6,5 [5,52; 12,88]	147/315 (46,7)	4,7 [3,68; 6,44]	0,81 [0,645; 1,024] 0,0783
Nein	71/140 (50,7)	8,2 [5,45; 11,93]	40/135 (29,6)	NE [11,30; NE]	1,70 [1,152; 2,504] 0,0068
Ätiologie HBV					

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[b] [95% KI]**	**n/N (%)**	**Median[b] [95% KI]**	**HR[c] [95% KI] p-Wert[d]**
Ja	107/244 (43,9)	11,9 [7,29; 13,83]	96/232 (41,4)	5,4 [4,53; 8,51]	0,82 [0,621; 1,082] 0,1602
Nein	109/201 (54,2)	5,5 [4,11; 7,43]	91/218 (41,7)	8,3 [5,45; 12,12]	1,25 [0,949; 1,656] 0,1105
BCLC Staging					
Stadium B	49/101 (48,5)	8,2 [6,57; 11,93]	27/88 (30,7)	NE [7,36; NE]	1,58 [0,986; 2,542] 0,0549
Stadium C	167/344 (48,5)	6,5 [5,52; 11,93]	160/362 (44,2)	5,5 [3,84; 7,46]	0,91 [0,734; 1,136] 0,4148
EORTC QLQ-C30 Kognitive Funktion					
Alle	301/445 (67,6)	2,8 [2,10; 3,65]	268/449 (59,7)	3,0 [2,76; 3,71]	1,09 [0,920; 1,291] 0,3167
Region[a]					
Asien-Pazifik	203/309 (65,7)	3,6 [2,73; 4,53]	187/312 (59,9)	2,8 [2,17; 3,65]	0,93 [0,764; 1,139] 0,4952
Westliche Regionen	98/136 (72,1)	1,9 [1,84; 2,79]	81/137 (59,1)	5,5 [2,79; 7,29]	1,49 [1,106; 1,997] 0,0082
EORTC QLQ-C30 Soziale Funktion					
Alle	293/445 (65,8)	2,8 [1,91; 3,02]	264/450 (58,7)	2,8 [2,30; 3,68]	1,06 [0,899; 1,260] 0,4658
Leberzirrhose zu Baseline (IIR)					
Ja	211/327 (64,5)	2,8 [1,91; 3,68]	213/342 (62,3)	2,7 [1,87; 3,42]	0,92 [0,762; 1,116] 0,4037
Nein	82/118 (69,5)	2,7 [1,87; 3,71]	51/108 (47,2)	4,6 [2,76; 15,31]	1,53 [1,078; 2,170] 0,0164
EORTC QLQ-C30 Globaler Gesundheitsstatus					
Alle	307/445 (69,0)	2,7 [1,87; 2,79]	278/450 (61,8)	2,1 [1,87; 2,79]	0,99 [0,841; 1,171] 0,9247
Baseline Gewichtsgruppe					

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[b] [95% KI]**	**n/N (%)**	**Median[b] [95% KI]**	**HR[c] [95% KI] p-Wert[d]**
<60 kg	90/145 (62,1)	2,8 [2,04; 5,62]	89/141 (63,1)	2,0 [1,84; 3,61]	0,77 [0,572; 1,037] 0,0840
≥60 kg	217/300 (72,3)	2,0 [1,87; 2,79]	189/309 (61,2)	2,7 [1,87; 2,79]	1,11 [0,910; 1,345] 0,3082
Ätiologie HBV					
Ja	159/244 (65,2)	3,7 [2,63; 4,63]	142/232 (61,2)	2,0 [1,87; 3,65]	0,83 [0,660; 1,042] 0,1068
Nein	148/201 (73,6)	1,9 [1,77; 2,69]	136/218 (62,4)	2,3 [1,87; 2,79]	1,22 [0,964; 1,536] 0,0986

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.

b: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

c: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

d: Der p-Wert wurde anhand eines unstratifizierten Log-Rank-Tests berechnet.

BCLC: Barcelona Clinic Liver Cancer; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; ES: Extrahepatische Ausbreitung (Extrahepatic Spread); HBV: Hepatitis B Virus; HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); ITT: Intention-to-treat; kg: Kilogramm; KI: Konfidenzintervall; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in der ITT-Population mit einem Baseline und mind. einem Folgewert; NE: not estimable (nicht berechenbar)

Quelle: SCO:SSiS 2018

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Physische Funktion, gemessen auf der Skala des EORTC QLQ-C30, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Ätiologie HBV“ festgestellt (Tabelle 4-65). Für Patienten mit HBV als Ursache für ihre Erkrankung ist der Effekt gleichgerichtet wie in der Gesamtpopulation. Für alle anderen Patienten zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant (Tabelle 4-66). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Emotionale Funktion, gemessen auf der Skala des EORTC QLQ-C30, wurden Belege für

Interaktion (Interaktions-p-Wert < 0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-65):

- Region
- MPVI
- MPVI oder ES oder beides
- Ätiologie HBV
- BCLC Staging

Für Patienten aus der Region Asien-Pazifik zeigt der Effekt in die entgegengesetzte Richtung wie der Gesamteffekt, ist aber nicht statistisch signifikant. Für Patienten aus den westlichen Regionen zeigt der Effekt in die gleiche Richtung wie der Effekt in der Gesamtpopulation. Für Patienten mit einer vorliegenden MPVI bzw. MPVI oder ES oder beides ist der Effekt anders gerichtet als in der Gesamtpopulation, aber nicht statistisch signifikant. Für Patienten ohne vorliegende MPVI bzw. MPVI oder ES oder beides zeigt der Effekt in die gleiche Richtung wie der Gesamteffekt. Hinsichtlich des Merkmals „Ätiologie HBV" zeigt der Effekt für Patienten mit HBV als Ursache für ihre Erkrankung in die entgegengesetzte Richtung wie der Effekt in der Gesamtpopulation, ist aber nicht statistische signifikant. Der Effekt für Patienten, deren Ursache für ihre Erkrankung nicht in HBV liegt, zeigt der Effekt in die gleiche Richtung wie der Gesamteffekt. Bei der Betrachtung des „BCLC Staging" zeigt der Effekt für Patienten mit BCLC-Stadium B in die gleiche Richtung wie in der Gesamtpopulation. Für Patienten mit BCLC-Stadium C zeigt der Effekt in die andere Richtung, ist aber nicht statistisch signifikant (Tabelle 4-66). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Kognitive Funktion, gemessen auf der Skala des EORTC QLQ-C30, wurde ebenfalls ein Beleg für Interaktion (Interaktions-p-Wert <0,05) bezüglich des Subgruppenmerkmals „Region" festgestellt (Tabelle 4-65). Für Patienten aus der Region Asien-Pazifik zeigt der Effekt in die entgegengesetzte Richtung wie der Gesamteffekt, ist aber nicht statistisch signifikant. Für Patienten aus den westlichen Regionen zeigt der Effekt in die gleiche Richtung wie der Effekt in der Gesamtpopulation (Tabelle 4-66). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Soziale Funktion, gemessen auf der Skala des EORTC QLQ-C30, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Leberzirrhose zu Baseline (IIR)" festgestellt (Tabelle 4-65). Für Patienten mit Leberzirrhose zu Baseline ist der Effekt anders gerichtet als in der Gesamtpopulation, ist aber nicht statistisch signifikant. Für alle anderen Patienten zeigt der Effekt in die gleiche Richtung (Tabelle 4-66).

Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Globaler Gesundheitsstatus, gemessen auf der Skala des EORTC QLQ-C30, wurden Belege für Interaktion (Interaktions-p-Wert <0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-65):

- Baseline Gewichtsgruppe
- Ätiologie HBV

Der Effekt für Patienten mit einem Körpergewicht kleiner 60 kg zeigt in die gleiche Richtung wie der Gesamteffekt. Für Patienten mit einem Körpergewicht über 60 kg zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant. Hinsichtlich des Merkmals „Ätiologie HBV“ zeigt der Effekt für Patienten mit HBV als Ursache für ihre Erkrankung in die gleiche Richtung wie der Effekt in der Gesamtpopulation. Der Effekt für Patienten, deren Ursache für ihre Erkrankung nicht in HBV liegt, zeigt der Effekt in die entgegengesetzte Richtung wie der Gesamteffekt, ist aber nicht statistisch signifikant (Tabelle 4-66). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.1.3.2.3.2 Subgruppenanalyse für die EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität)

Tabelle 4-67: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	**Interaktions-p-Wert[a]**		
	Körperbild	**Ernährung**	**Sexualleben**
Baseline Altersgruppe,	0,5864	0,5364	0,1644
Geschlecht	0,0941	**0,0320**	0,0908
Region	0,6472	0,7935	0,3490
MPVI	0,1891	0,2233	0,1141
ES	0,8669	0,7160	0,9910
MPVI oder ES oder beides	0,2005	**0,0199**	0,2499
Baseline ECOG-PS	0,9032	0,2954	0,0700
Baseline Gewichtsgruppe	0,8870	0,4197	0,8575
Baseline AFP	**0,0310**	0,0919	0,8486
Ätiologie HBV	0,0809	0,3069	0,5429

Subgruppen-Kategorie	Interaktions-p-Wert[a]		
	Körperbild	**Ernährung**	**Sexualleben**
Ätiologie HCV	0,6014	0,6632	0,7366
Ätiologie Alkohol	0,9314	0,9666	0,5531
BCLC Staging	0,3569	0,2884	0,6260
Leberzirrhose zu Baseline (CRF)	0,3015	0,8108	0,4701
Leberzirrhose zu Baseline (IIR)	0,2178	0,0904	**0,0169**
Vorbehandlung mit TACE	0,0808	0,9768	0,7437

a. Der Interaktions-p-Wert wurde anhand eines zweiseitigen unstratifizierten Cox-Proportional-Hazard-Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); TACE: Transarterielle Chemoembolisation

Quelle: SCO:SSiS 2018

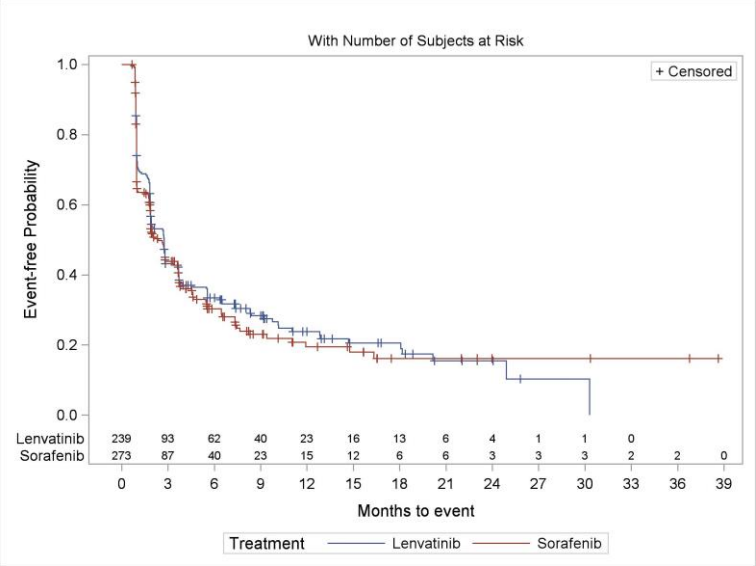


Abbildung 4-127: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 (Baseline AFP < 200 ng/mL) aus Studie 304 mit dem zu bewertenden Arzneimittel

AFP: Alpha-Fetoprotein; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; mL: Milliliter; ng: Nanogramm
Quelle: SCO:SSiS 2018

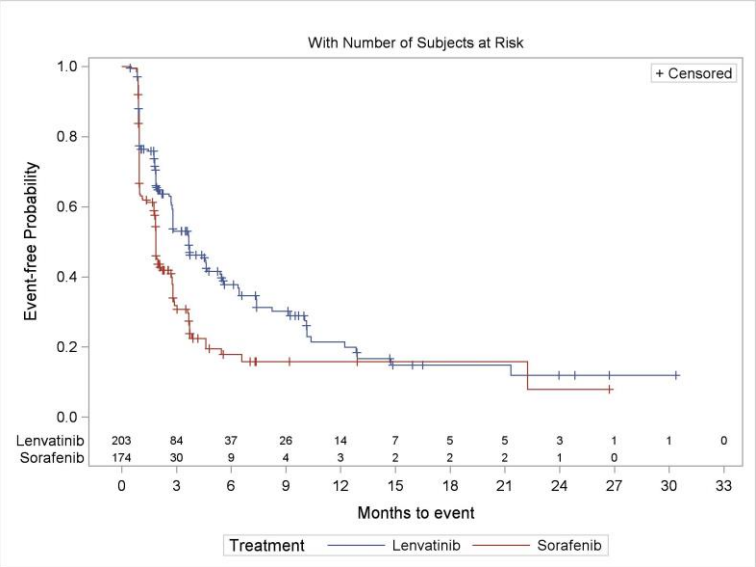


Abbildung 4-128: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Körperbild** des EORTC QLQ-HCC18 (Baseline AFP ≥ 200 ng/mL) aus Studie 304 mit dem zu bewertenden Arzneimittel

AFP: Alpha-Fetoprotein; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; mL: Milliliter; ng: Nanogramm
Quelle: SCO:SSiS 2018

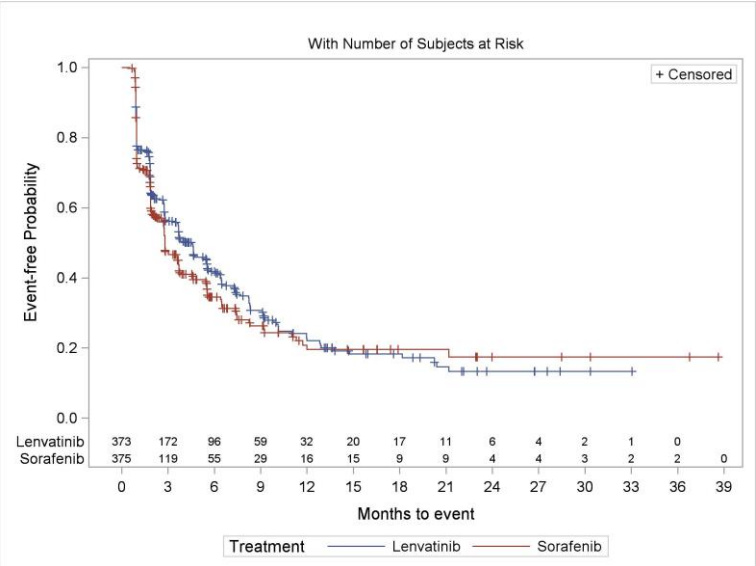


Abbildung 4-129: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (Geschlecht „Männlich“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

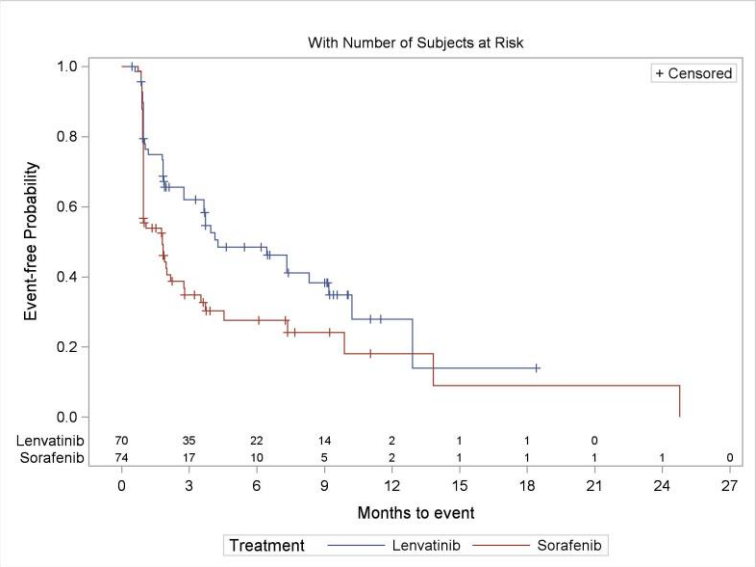


Abbildung 4-130: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (Geschlecht „Weiblich“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire
Quelle: SCO:SSiS 2018

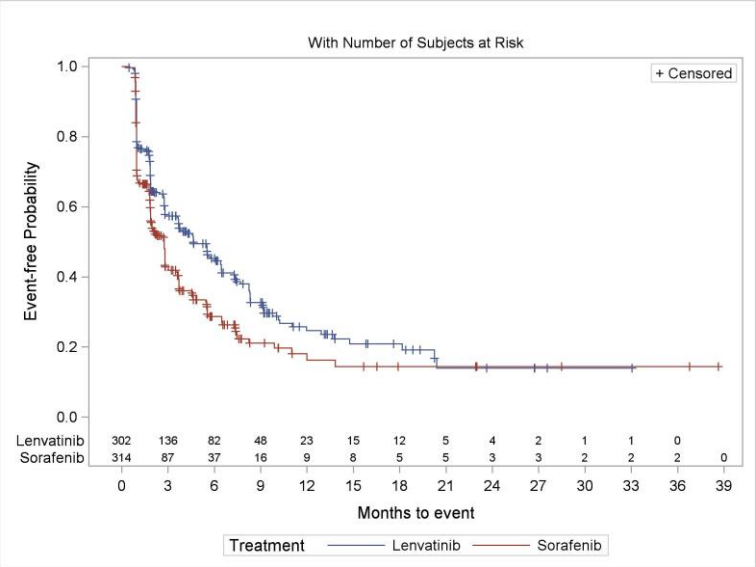


Abbildung 4-131: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (MPVI oder ES oder beides „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

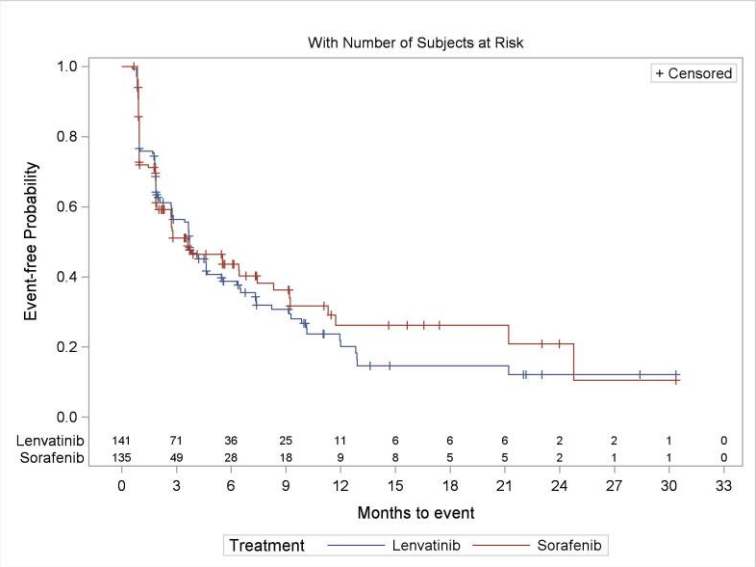


Abbildung 4-132: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Ernährung** des EORTC QLQ-HCC18 (MPVI oder ES oder beides „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader)
Quelle: SCO:SSiS 2018

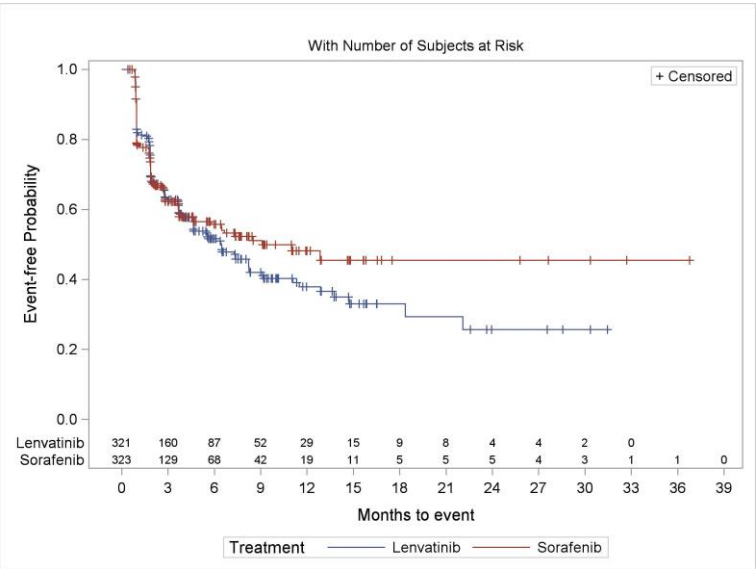


Abbildung 4-133: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 (Leberzirrhose zu Baseline (IIR) „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; IIR: Independent Imaging Review (unabhängige radiologische Bewertung)
Quelle: SCO:SSiS 2018

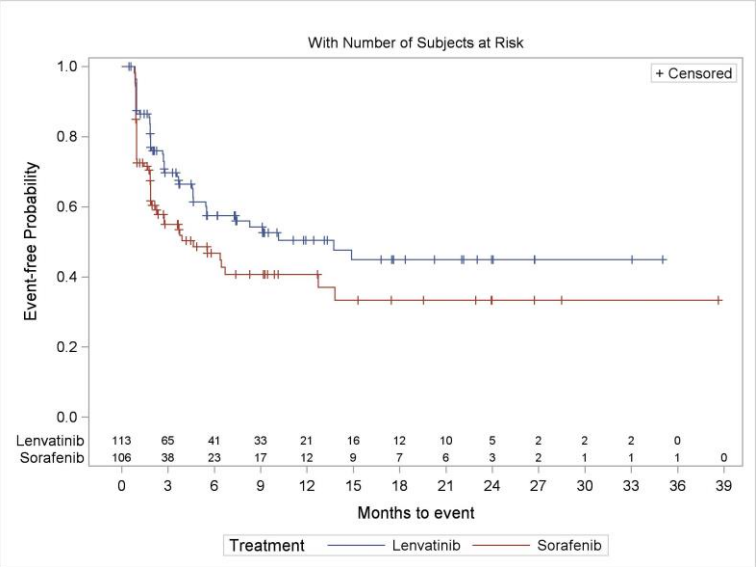


Abbildung 4-134: Subgruppenanalyse: Kaplan-Meier-Kurve für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf der Skala **Sexualleben** des EORTC QLQ-HCC18 (Leberzirrhose zu Baseline (IIR) „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; IIR: Independent Imaging Review (unabhängige radiologische Bewertung)
Quelle: SCO:SSiS 2018

Tabelle 4-68: Subgruppenanalyse: Ergebnisse für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score gemessen auf den Skalen des EORTC QLQ-HCC18 (gesundheitsbezogene Lebensqualität) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95% KI]**	**n/N (%)**	**Median[a] [95% KI]**	**HR[b] [95% KI] p-Wert[c]**
EORTC QLQ-HCC18 Körperbild					
Alle	299/443 (67,5)	2,8 [2,73; 3,68]	297/449 (66,1)	1,9 [1,87; 2,73]	0,79 [0,671; 0,932] 0,0051
Baseline AFP					
< 200 ng/mL	173/239 (72,4)	2,8 [1,91; 2,83]	179/273 (65,6)	2,4 [1,87; 2,83]	0,93 [0,753; 1,146] 0,4930
≥ 200 ng/mL	125/203 (61,6)	3,7 [2,79; 4,63]	117/174 (67,2)	1,9 [1,81; 2,17]	0,61 [0,469; 0,787] 0,0001

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[a] [95% KI]**	**n/N (%)**	**Median[a] [95% KI]**	**HR[b] [95% KI] p-Wert[c]**
EORTC QLQ-HCC18 Ernährung					
Alle	279/443 (63,0)	4,3 [3,68; 5,52]	274/449 (61,0)	2,8 [2,07; 3,02]	0,80 [0,674; 0,949] 0,0102
Geschlecht					
Männlich	240/373 (64,3)	4,6 [3,65; 5,52]	222/375 (59,2)	2,8 [2,73; 3,68]	0,88 [0,731; 1,055] 0,1655
Weiblich	39/70 (55,7)	4,3 [2,76; 9,20]	52/74 (70,3)	1,8 [0,95; 2,76]	0,58 [0,381; 0,888] 0,0112
MPVI oder ES oder beides					
Ja	180/302 (59,6)	4,6 [3,68; 6,41]	198/314 (63,1)	2,8 [1,91; 2,79]	0,71 [0,575; 0,866] 0,0008
Nein	99/141 (70,2)	3,7 [2,76; 4,63]	76/135 (56,3)	3,6 [2,69; 6,44]	1,09 [0,808; 1,471] 0,5706
EORTC QLQ-HCC18 Sexualleben					
Alle	206/434 (47,5)	7,4 [5,49; 9,20]	183/429 (42,7)	6,7 [4,63; 13,80]	0,94 [0,771; 1,154] 0,5700
Leberzirrhose zu Baseline (IIR)					
Ja	158/321 (49,2)	6,4 [4,63; 8,25]	129/323 (39,9)	9,1 [5,82; NE]	1,11 [0,878; 1,399] 0,3864
Nein	48/113 (42,5)	13,7 [5,45; NE]	54/106 (50,9)	4,6 [2,17; 12,71]	0,64 [0,436; 0,953] 0,0264

a: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt.

b: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet.

c: Der p-Wert wurde anhand eines unstratifizierten Log-Rank-Tests berechnet.

AFP: Alpha-Fetoprotein; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); ITT: Intention-to-treat; KI: Konfidenzintervall; mL: Milliliter; MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); n: Anzahl der Patienten mit einem Ereignis; N: Anzahl der Patienten in dr ITT-Population mit einem Baseline- und mind. einem Folgewert; NE: not estimable (nicht berechenbar); ng: Nanogramm

Quelle: SCO:SSiS 2018

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Körperbild, gemessen auf der Skala des EORTC QLQ-HCC18, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Baseline AFP“ festgestellt (Tabelle 4-67). Für Patienten mit einem Wert kleiner 200 ng/mL sowie für Patienten mit einem Wert größer bzw. gleich 200 ng/mL ist der Effekt gleichgerichtet wie in der Gesamtpopulation (Tabelle 4-68). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Ernährung, gemessen auf der Skala des EORTC QLQ-HCC18, wurden Belege für Interaktion (Interaktions-p-Wert <0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-67):

- Geschlecht
- MPVI oder ES oder beides

Der Effekt für weibliche als auch für männliche Patienten zeigt in die gleiche Richtung wie der Gesamteffekt. Für Patienten, bei denen eine MPVI oder ES oder beides vorlag, ist der Effekt ebenfalls gleichgerichtet wie in der Gesamtpopulation. Für Patienten, bei denen keine MPVI oder ES oder beides vorlag, ist der Effekt entgegengesetzt gerichtet, aber nicht statistisch signifikant (Tabelle 4-68). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

In der Studie 304 wurde für die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score des Items Sexualleben, gemessen auf der Skala des EORTC QLQ-HCC18, ein Beleg für Interaktion (Interaktions-p-Wert < 0,05) bezüglich des Subgruppenmerkmals „Leberzirrhose zu Baseline (IIR)“ festgestellt (Tabelle 4-67). Für Patienten mit Leberzirrhose zu Baseline ist der Effekt anders gerichtet als in der Gesamtpopulation, aber nicht statistisch signifikant. Für Patienten ohne Leberzirrhose zu Baseline hingegen ist der Effekt gleichgerichtet (Tabelle 4-68). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.1.3.2.4 Subgruppenanalyse für die Endpunkte zu Sicherheit und Verträglichkeit

Tabelle 4-69: Ergebnisse der Tests auf Interaktion der Subgruppenanalysen für die Zeit bis zum Auftreten des ersten UE (in Tagen) aus Studie 304 mit dem zu bewertenden Arzneimittel

Subgruppen-Kategorie	Interaktions-p-Wert[a]			
	UE	**SUE**	**UE, die zum Therapieabbruch führten**	**UE mit CTCAE Grad ≥ 3**
Baseline Altersgruppe,	0,3737	0,2266	0,7773	0,6361
Geschlecht	0,5919	0,4947	0,7797	0,2851
Region	0,1481	0,6305	0,3781	**0,0094**
MPVI	0,1670	0,9866	0,8195	0,0637
ES	0,6241	0,6456	0,9640	0,7880
MPVI oder ES oder beides	0,6330	0,2961	0,9362	0,1751
Baseline ECOG-PS	0,6140	0,1921	0,7422	0,7390
Baseline Gewichtsgruppe	0,8605	0,0676	0,4216	**0,0002**
Baseline AFP	0,3075	0,5023	0,7189	0,4019
Ätiologie HBV	**0,0294**	0,3122	0,9210	0,9321
Ätiologie HCV	**0,0173**	0,0897	0,7684	0,5097
Ätiologie Alkohol	0,6780	0,8875	0,5119	0,8081
BCLC Staging	0,6784	0,1540	0,8997	0,1042
Leberzirrhose zu Baseline (CRF)	0,6989	0,3816	0,7521	0,6535
Leberzirrhose zu Baseline (IIR)	0,9306	0,6244	0,9929	0,2812
Vorbehandlung mit TACE	0,0915	0,4030	**0,0454**	0,5999

a. Der Interaktions-p-Wert wurde anhand eines generalisierten linearen Modells, mit „Behandlung“, „Subgruppenmerkmal“ und „Interaktion zwischen Behandlung und Subgruppenmerkmal“ als Faktor, ermittelt.

AFP: Alpha-Fetoprotein; BCLC: Barcelona Clinic Liver Cancer; CRF: Case Report Form; CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); ECOG; ECOG-PS: Eastern Cooperative Oncology Group-Performance Status; ES: Extrahepatic Spread (Extrahepatische Ausbreitung); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); MPVI: Macroscopic Portal Vein Invasion (Makroskopische Invasion der Pfortader); SUE: Schwerwiegendes UE; TACE: Transarterielle Chemoembolisation; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

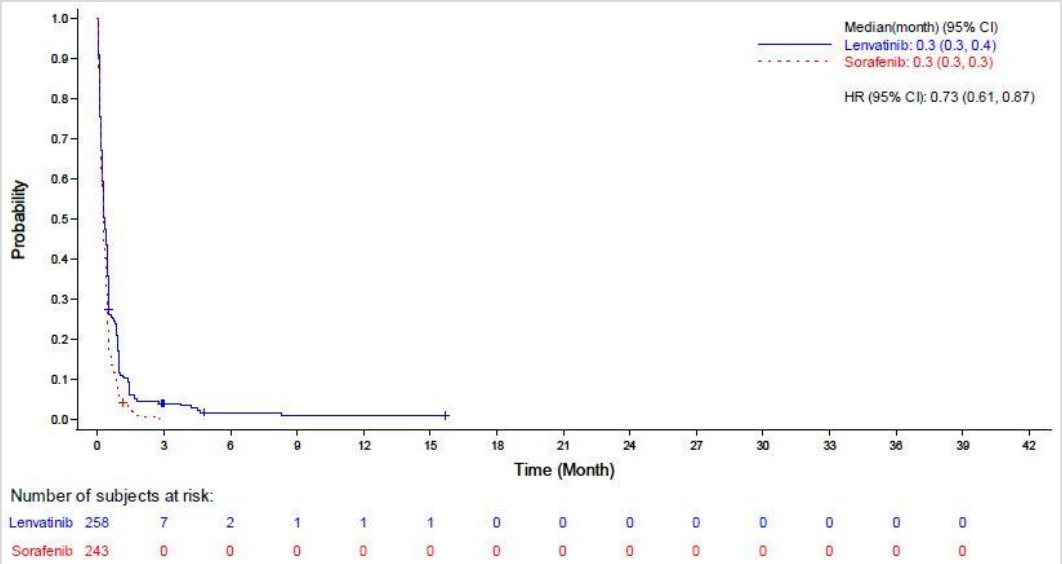


Abbildung 4-135: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HBV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel
CI: Confidence Interval (Konfidenzintervall); HBV: Hepatitis B Virus; HR: Hazard Ratio; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

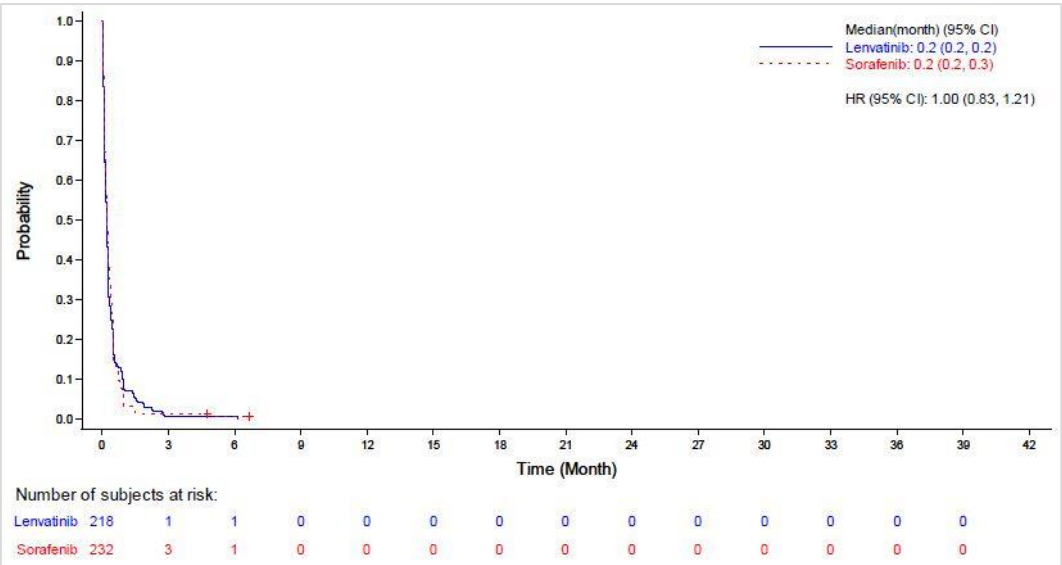


Abbildung 4-136: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HBV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel
CI: Confidence Interval (Konfidenzintervall); HBV: Hepatitis B Virus; HR: Hazard Ratio; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

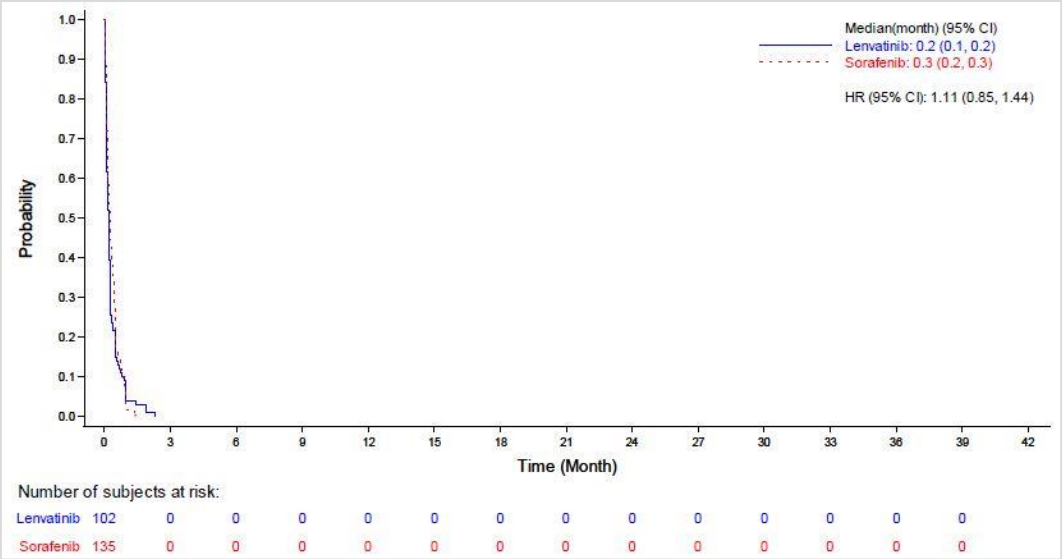


Abbildung 4-137: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HCV „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel
CI: Confidence Interval (Konfidenzintervall); HCV: Hepatitis C Virus; HR: Hazard Ratio; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

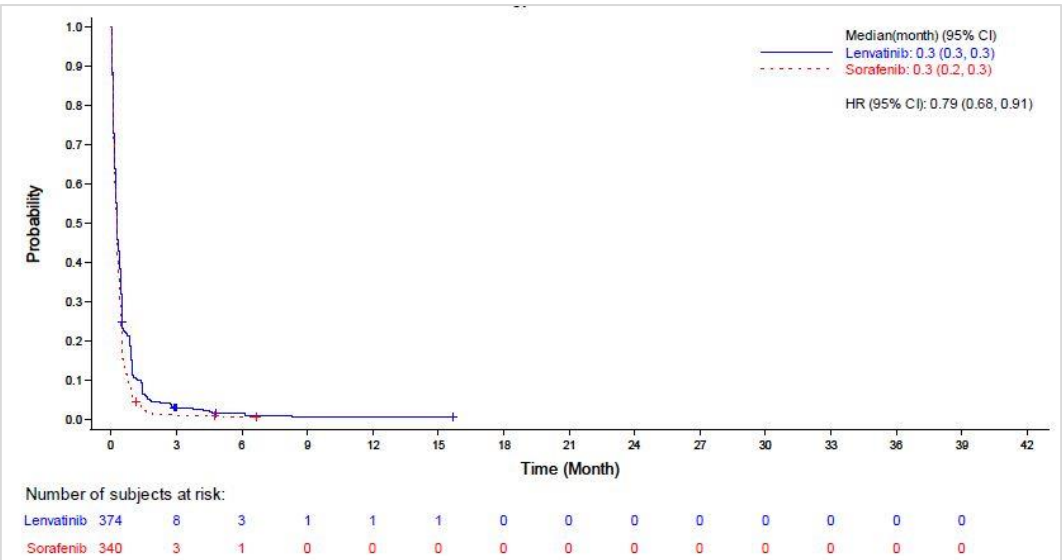


Abbildung 4-138: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE** (Ätiologie HCV „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel
CI: Confidence Interval (Konfidenzintervall); HCV: Hepatitis C Virus; HR: Hazard Ratio; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

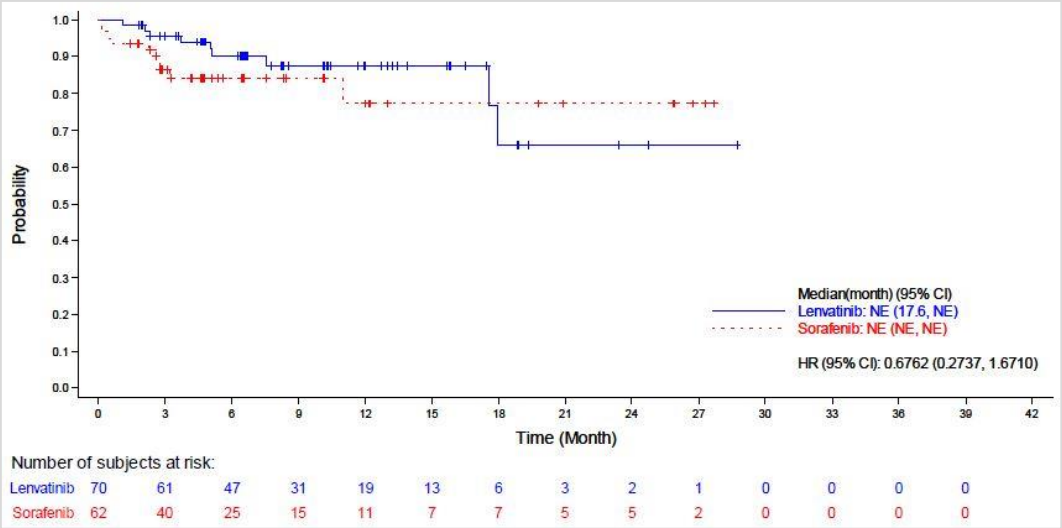


Abbildung 4-139: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE, die zum Therapieabbruch führten** (Vorbehandlung mit TACE „Ja“) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; TACE: Transarterielle Chemoembolisation; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

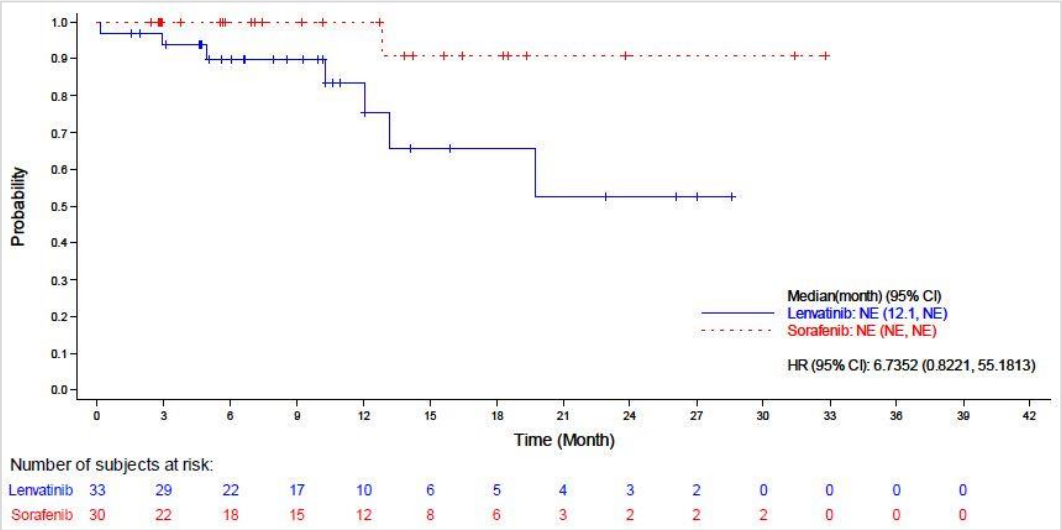


Abbildung 4-140: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE, die zum Therapieabbruch führten** (Vorbehandlung mit TACE „Nein“) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); HR: Hazard Ratio; TACE: Transarterielle Chemoembolisation; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

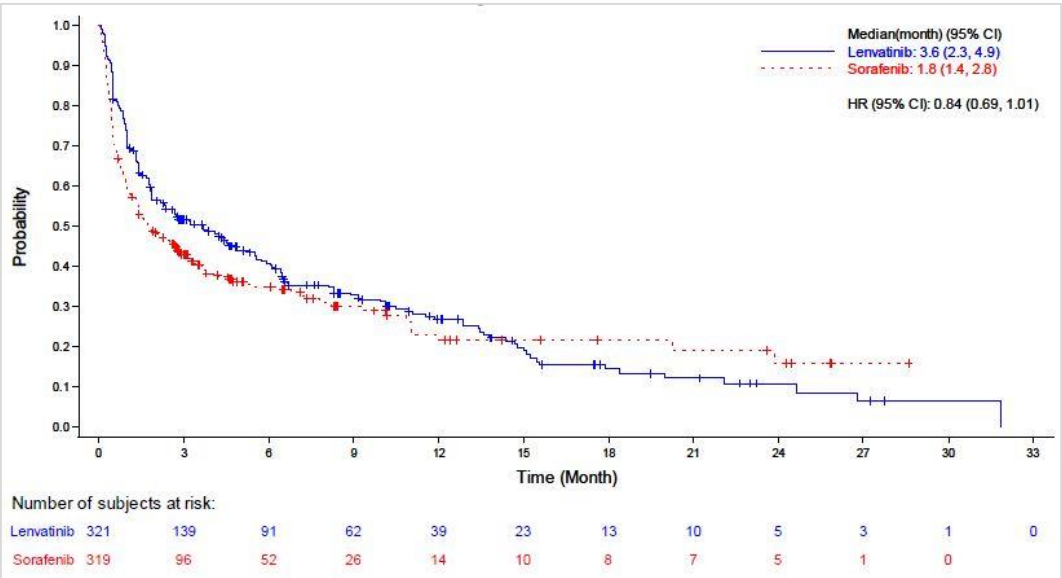


Abbildung 4-141: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Region[a] „Asien-Pazifik“) aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

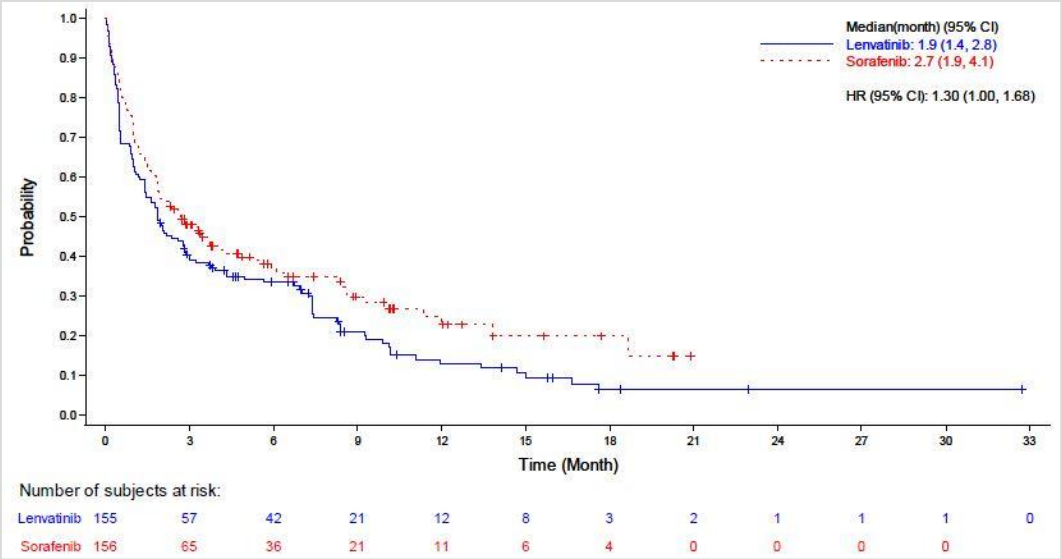


Abbildung 4-142: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Region[a] „Westliche Regionen“) aus Studie 304 mit dem zu bewertenden Arzneimittel

a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand.
CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse
Quelle: Eisai Inc. 2018c

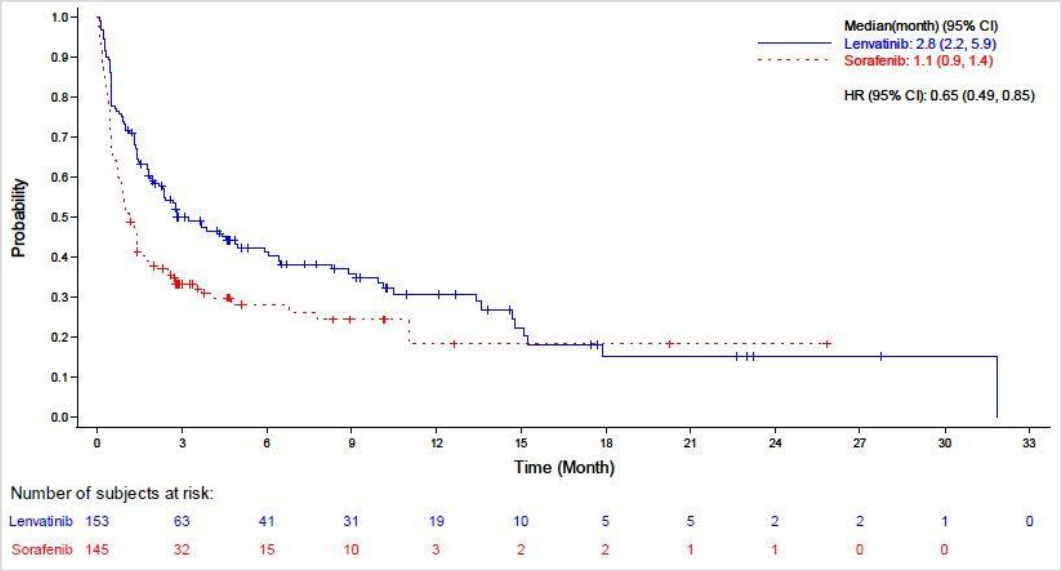


Abbildung 4-143: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Baseline Gewichtsgruppe „< 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; kg: Kilogramm; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

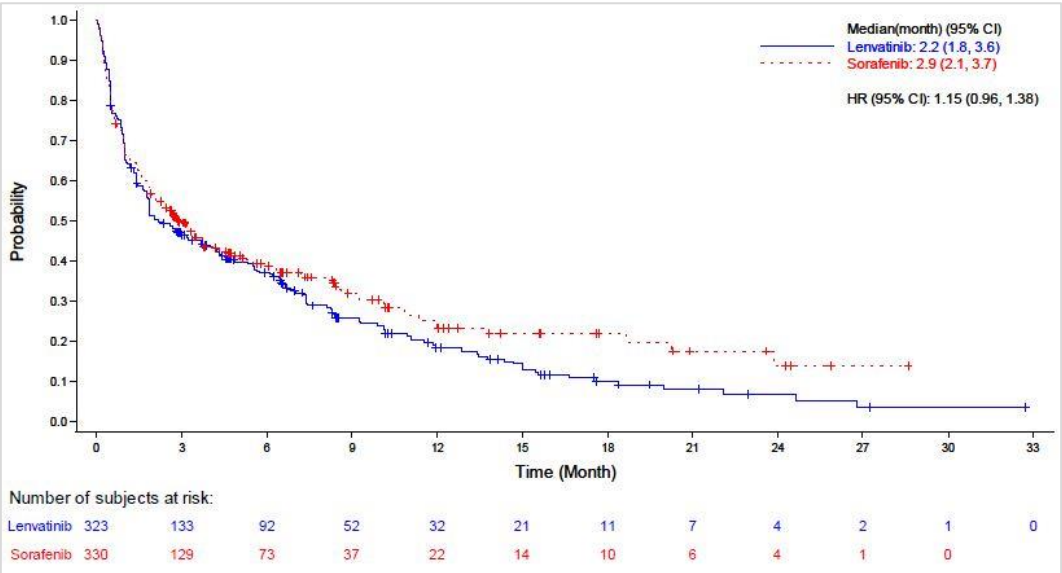


Abbildung 4-144: Subgruppenanalyse: Kaplan-Meier-Kurve für den Endpunkt **UE mit CTCAE Grad ≥ 3** (Baseline Gewichtsgruppe „≥ 60 kg“) aus Studie 304 mit dem zu bewertenden Arzneimittel

CI: Confidence Interval (Konfidenzintervall); CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HR: Hazard Ratio; kg: Kilogramm; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse

Quelle: Eisai Inc. 2018c

Tabelle 4-70: Subgruppenanalyse: Ergebnisse für die Zeit bis zum Auftreten des ersten UE (in Tagen) stratifiziert nach den Subgruppenmerkmalen mit einem Beleg für Interaktion aus Studie 304 mit dem zu bewertenden Arzneimittel

	Lenvatinib		**Sorafenib**		**Lenvatinib vs. Sorafenib**
Subgruppen-Kategorie	**n/N (%)**	**Median[b] [95% KI]**	**n/N (%)**	**Median[b] [95% KI]**	**HR[c] [95% KI]**
UE					
Alle	470/476 (98,7)	8 [7; 8]	472/475	8 [7; 8]	0,84 [0,74; 0,95]
Ätiologie HBV					
Ja	252/258 (97,7)	10 [8; 12]	242/243	8 [8; 10]	0,73 [0,61; 0,87]
Nein	218/218 (100,0)	6 [5; 7]	230/232	7 [6; 8]	1,00 [0,83; 1,21]
Ätiologie HCV					
Ja	102/102 (100,0)	6 [4; 7]	135/135	8 [6; 10]	1,11 [0,85; 1,44]
Nein	368/374 (95,4)	8 [8; 10]	337/340	8 [7; 9]	0,79 [0,68; 0,91]
UE, die zum Therapieabbruch führten					
Alle	94/476 (75,0)	82 [57; 112]	316/475 (66,5)	64 [50; 86]	0,96 [0,83; 1,12]
Vorbehandlung mit TACE					
Ja	9/70 (12,9)	NE [535, NE]	10/62 (16,1)	NE [NE; NE]	0,68 [0,27; 1,67]
Nein	7/33 (21,2)	NE [367; NE]	1/30 (3,3)	NE [NE; NE]	6,74 [0,82; 55,18]
UE mit CTCAE Grad ≥ 3					
Alle	357/476 (75,0)	82 [57; 112]	316/475 (66,5)	64 [50; 86]	0,96 [0,83; 1,12]
Region[a]					
Asien-Pazifik	228/321 (71,0)	111 [69, 149]	210/319 (65,8)	55 [42; 84]	0,84 [0,69; 1,01]
Westliche Regionen	129/155 (83,2)	57 [42; 85]	106/156 (67,9)	82 [57; 125]	1.30 [1,00; 1.68]
Baseline Gewichtsgruppe					
< 60 kg	102/153 (66,7)	86 [67; 180]	103/145 (71,0)	34 [26; 43]	0,65 [0,49; 0,85]
≥ 60 kg	255/323 (78,9)	66 [55; 111]	213/330 (64,5)	87 [64; 113]	1,15 [0,96; 1,38]

	Lenvatinib		Sorafenib		Lenvatinib vs. Sorafenib
Subgruppen-Kategorie	**n/N (%)**	**Median[b] [95% KI]**	**n/N (%)**	**Median[b] [95% KI]**	**HR[c] [95% KI]**
a: Die westlichen Regionen bestehen aus Nordamerika und Europa, inklusive Russland und Israel; die asiatisch-pazifischen Regionen bestehen aus China, Hong Kong, Japan, Korea, Malaysia, Philippinen, Singapur, Taiwan und Thailand. b: Der Median wurde basierend auf der Kaplan-Meier-Methode ermittelt. Das zweiseitige 95%-KI wurde mithilfe der generalisierten Brookmeyer und Crowley Methode bestimmt. c: Die HR der Subgruppenausprägungen wurde anhand eines unstratifizierten Cox-Proportional-Hazard-Modells berechnet. CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); HBV: Hepatitis B Virus; HCV: Hepatitis C Virus; HR: Hazard Ratio; kg: Kilogramm; KI: Konfidenzintervall; n: Anzahl der Patienten mit Ereignis; N: Anzahl der Patienten in der SAS-Population; NE: not estimable (nicht berechenbar); SAS: Safety Analysis Set; TACE: Transarterielle Chemoembolisation; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse Quelle: Eisai Inc. 2018c					

Hinsichtlich der Zeit bis zum Auftreten des ersten UE wurden Belege für Interaktion (Interaktions-p-Wert <0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-69):

- Ätiologie HBV
- Ätiologie HCV

Für Patienten mit HBV als Ursache für ihre Erkrankung ist der Effekt gleichgerichtet wie in der Gesamtpopulation. Für alle anderen Patienten zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant. Für Patienten mit HCV als Ursache für ihre Erkrankung ist der Effekt anders gerichtet als in der Gesamtpopulation, aber nicht statistisch signifikant. Für alle anderen Patienten zeigt der Effekt in die gleiche Richtung (Tabelle 4-70). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zum Auftreten des ersten UE, das zum Therapieabbruch führte, wurde ein Beleg für Interaktion (Interaktions-p-Wert <0,05) bezüglich des Subgruppenmerkmals „Vorbehandlung mit TACE“ festgestellt (Tabelle 4-69). Für Patienten mit einer TACE Vorbehandlung ist der Effekt gleichgerichtet wie in der Gesamtpopulation. Für alle anderen Patienten zeigt der Effekt in die andere Richtung, aber nicht statistisch signifikant (Tabelle 4-70). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

Hinsichtlich der Zeit bis zum Auftreten des ersten UE mit CTCAE Grad ≥ 3, wurden Belege für Interaktion (Interaktions-p-Wert <0,05) bezüglich der folgenden Subgruppenmerkmale festgestellt (Tabelle 4-69):

- Region
- Baseline Gewichtsgruppe

Der Effekt für Patienten aus der Asien-Pazifik Region zeigt in die gleiche Richtung wie der Gesamteffekt. Für Patienten aus den westlichen Regionen zeigt der Effekt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant, da sich die Konfidenzintervalle der HR für westliche Regionen und der Asien-Pazifik Region überschneiden. Für Patienten mit einem Baseline-Gewicht kleiner 60 kg zeigt der Effekt ebenfalls in die gleiche Richtung wie der Gesamteffekt. Der Effekt für Patienten über 60 kg zu Baseline zeigt in die entgegengesetzte Richtung, ist aber nicht statistisch signifikant (Tabelle 4-70). Somit konnte kein für die Nutzenbewertung relevanter Effektmodifikator identifiziert werden. Die Ableitung des Zusatznutzens erfolgt auf Basis der Gesamtpopulation.

4.3.2 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Zur Bewertung des medizinischen Zusatznutzens von Lenvatinib als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben, werden die Ergebnisse einer multizentrischen, randomisierten, aktivkontrollierten und Open-Label durchgeführten Nicht-Unterlegenheits- Phase-III-Studie, der Studie 304, zu den Endpunkten zu Mortalität, Morbidität, gesundheitsbezogene Lebensqualität sowie Sicherheit und Verträglichkeit herangezogen. Das Verzerrungspotenzial der Studie 304 ist insgesamt als niedrig einzustufen.

Mortalität

Für den Wirksamkeitsendpunkt OS zeigte sich in der betrachteten Studienpopulation bezüglich des HR ein numerischer Vorteil für Lenvatinib gegenüber Sorafenib basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell (HR=0,92; 95%-KI [0,79; 1,06]; p=0,2526). Da die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag, konnte hiermit die Nicht-Unterlegenheit von Lenvatinib gegenüber Sorafenib belegt werden, was einen Überlebensvorteil von Lenvatinib gegenüber Placebo impliziert. Die mediane Überlebensdauer unter Lenvatinib betrug 13,6 Monate (95%-KI [12,1; 14,9]), unter Sorafenib betrug sie 12,3 Monate (95%-KI [10,4; 13,9]). Damit ergibt sich ein absoluter Vorteil im medianen Überleben von 1,3 Monaten zugunsten von Lenvatinib.

Die Studie 304 ist somit die erste positive, globale Phase-III-Studie in der Erstlinientherapie des HCC in den letzten 10 Jahren. Darüber hinaus ist die Studie 304 die erste positive Studie

im HCC mit einer aktiven Kontrollgruppe. Seit der Zulassung von Sorafenib gab es vier Phase-III-Studien mit zielgerichteten Wirkstoffen, die keine Überlegenheit oder Nicht-Unterlegenheit in der Erstlinientherapie zeigen konnten (Kudo et al. 2018).

Die Ergebnisse der Sensitivitätsanalysen, die auf uni- und multivariater Adjustierung basierten, zeigten konsistent gleichgerichtete Effekte wie die Ergebnisse der Hauptanalyse und bestätigen somit deren Robustheit. In Bezug auf die AFP adjustierte und multivariate adjustierte Analyse zeigte Lenvatinib im Vergleich zu Sorafenib sogar eine statistisch signifikante Überlegenheit im OS. Die in den adjustierten Analysen nachgewiesene statistische Signfikanz kann auf die Imbalance in den Behandlungsarmen bezüglich des für das HCC prognostischen Faktors AFP zurückgeführt werden.

Zusammenfassend liegen für Lenvatinib gegenüber Sorafenib eine Nicht-Unterlegenheit und ein numerischer Vorteil bezüglich der Hauptanalyse vor. Dieser numerische Vorteil im OS wird nach Adjustierung für prognostische Faktoren statistisch signifikant und weist damit auf eine Überlegenheit im OS hin.

Morbidität

In der Hauptanalyse des Endpunkts PFS auf Basis des IIR per mRECIST zeigte sich im HR eine statistisch signifikante und klinisch relevante Überlegenheit bezüglich der Risikoreduktion einer Progression von Lenvatinib gegenüber Sorafenib basierend auf dem Cox-Proportional-Hazard-Modell (HR 0,64; 95%-KI [0,55; 0,75]; p<0,0001). Unter der Behandlung mit Lenvatinib wurde somit eine 36%-ige Verringerung des Progressionsrisikos im Vergleich zur Behandlung mit Sorafenib erreicht. In Bezug auf die mediane Zeit bis zur Krankheitsprogression war das mediane PFS im Lenvatinibarm mit 7,3 Monaten (95%-KI [5,6; 7,5]) mehr als doppelt so lang wie im Sorafenibarm (3,6 Monaten (95%-KI [3,6; 3,7])). Somit ergab sich ein klinisch relevanter absoluter Vorteil von 3,7 Monaten für das mediane PFS zugunsten von Lenvatinib. Die Überlegenheit von Lenvatinib beim PFS konnte dabei unabhängig von der verwendeten Analysemethode bestätigt werden, wie durch die Ergebnisse der zwei durchgeführten Sensitivitätsanalysen gemäß IIR per RECIST v1.1 (HR 0,65; 95%-KI [0,56; 0,77]; p<0,0001) und Prüfarzt per mRECIST (HR 0,66; 95%-KI [0,57; 0,77]; p<0,0001) gezeigt wurde.

Die Ergebnisse der Analysen für den Endpunkt ORR zeigen sowohl in der Tumorbewertung gemäß IIR per mRECIST (OR=5,01; 95%-KI [3,59; 7,01]; p<0,0001) und RECIST v.1.1 (OR=3,34; 95%-KI [2,17; 5,14]; p<0,0001) als auch in der Tumorbewertung durch den Prüfarzt per mRECIST (OR=3,13; 95%-KI [2,15; 4,56]; p<0,0001) eine statistisch signifikante und klinisch relevante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die Zahl der Patienten, die auf eine Therapie mit Lenvatinib bezogen auf das Tumorwachstum ansprachen, war somit um mehr als das Dreifache (gemäß IIR per RECIST 1.1 und Prüfarzt per mRECIST) bzw. sogar mehr als das Fünffache (gemäß IIR per mRECIST) gegenüber Sorafenib erhöht. Legt man die Ergebnisse der unabhängigen radiologischen Untersuchung auf Grundlage der mRECIST Kriterien zugrunde, so heißt dies konkret, dass unter der Behandlung mit Lenvatinib mehr als jeder dritte Patient (40.6%) von der Therapie profitiert, verglichen mit weniger als jedem achten

Patienten (12,4%) unter der Behandlung mit Sorafenib. Zudem wurde bei 10 Patienten im Lenvatinibarm eine Complete Response (per mRECIST IIR) – das heißt ein Verschwinden aller Läsionen – beobachtet verglichen mit 4 Patienten im Sorafenibarm (Eisai Inc. 2017a). Die Ergebnisse für den Endpunkt ORR stützen damit die Ergebnisse des Endpunkts PFS.

Die Analyse der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur Morbidität des Fragebogens EORTC QLQ-C30 zeigte für die den Patienten beeinträchtigenden und belastenden Symptome Schmerz (HR=0,81; 95%-KI [0,690; 0,949]; p=0,0091) und Diarrhö (HR=0,52; 95%-KI [0,440; 0,623]; p<0,0001) eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf der Skala Schmerz betrug dabei 1,9 Monate (95%-KI [1,87; 2,76]) für Lenvatinib im Vergleich zu 1,8 Monate (95%-KI [1,12; 1,87]) für Sorafenib. Bezüglich der Skala Diarrhö betrug der absolute Vorteil 1,9 Monate in der medianen Zeit bis zur ersten Verschlechterung um 10 Punkte zugunsten von Lenvatinib (4,6 Monate; 95%-KI [3,71; 5,52]) im Vergleich zu Sorafenib (2,7 Monate; 95%-KI [1,87; 2,79]). Hinsichtlich der Symptome Fatigue, Übelkeit und Erbrechen, Dyspnoe, Schlaflosigkeit, Appetitminderung und Obstipation zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Auch im Hinblick auf die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur Morbidität des Fragebogens EORTC QLQ-HCC18 konnten die Patienten unter der Behandlung mit Lenvatinib für alle Symptome vergleichbare Therapieerfolge erzielen wie unter der Behandlung mit Sorafenib.

Die Analyse des Fragebogens EQ-5D bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS zeigte für die Patienten unter der Behandlung mit Lenvatinib ebenfalls vergleichbare Therapieerfolge wie unter Sorafenib.

Zusammenfassend besteht für Lenvatinib gegenüber Sorafenib eine statistisch signifikante und klinisch relevante Überlegenheit im PFS und ORR. Darüber hinaus konnte bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im Score der Symptome, gemessen auf den Skalen zur Morbidität des EORTC QLQ-C30 und -HCC18, eine statistisch signifikante Überlegenheit bei einzelnen Symptomen sowie ein vergleichbarer Therapieerfolg bei allen weiteren Symptomen von Lenvatinib gegenüber Sorafenib gezeigt werden. Auch bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS konnten vergleichbare Therapieerfolge erzielt werden.

Gesundheitsbezogene Lebensqualität

Bei der Analyse des Rollenfunktionsniveaus des Fragebogens EORTC QLQ-C30 zeigte sich bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf der entsprechenden Skala eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,82; 95%-KI [0,696; 0,962]; p=0,0148). Die mediane Zeit betrug dabei 2,0 Monate (95 5-KI [1,87; 2,83]) für Lenvatinib im Vergleich zu 1,9 Monate (95%-KI [1,84; 1,97]) für Sorafenib. Hinsichtlich der physischen, emotionalen, kognitiven, sozialen und finanziellen Funktion und des globalen Gesundheitsstatus zeigten sich unter der

Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib.

Auch die Analysen des Fragenbogens EORTC QLQ-HCC18 bezüglich Körperbild (HR=0,79; 95%-KI [0,671; 0,932]; p=0,0051) und Ernährung (HR=0,80; 95%-KI [0,674; 0,949]; p=0,0102) zeigten eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den entsprechenden Skalen. Bezüglich des Items Körperbild betrug die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte 2,8 Monate (95%-KI [2,73; 3,68]) für Lenvatinib im Vergleich zu 1,9 Monate (95 5-KI [1,87; 2,73]) für Sorafenib. Bezüglich des Items Ernährung betrug die Zeit 4,3 Monate (95%-KI [3,68; 5,52]) für Lenvatinib im Vergleich zu 2,8 Monate (95%-KI [2,07; 3,02]) für Sorafenib.

Zusammenfassend zeigte Lenvatinib gegenüber Sorafenib eine statistisch signifikante Überlegenheit bei drei Endpunkten der gesundheitsbezogenen Lebensqualität (gemessen als Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur gesundheitsbezogenen Lebensqualität des EORTC QLQ-C30 und -HCC18). Für die anderen Endpunkte wurde ein vergleichbarer Therapieerfolg nachgewiesen.

Sicherheit und Verträglichkeit

Da die Gesamtbeobachtungszeiten im Behandlungsarm Lenvatinib um mehr als 75 Tage länger waren als in der Kontrollgruppe, wären die Inzidenzraten und die darauf basierenden Effektmaße RR, OR und RD zuungunsten von Lenvatinib verzerrt, so dass nur sehr eingeschränkt Aussagen zu einem größeren oder vergleichbaren bzw. geringeren Schaden gemacht werden könnten. Aus diesem Grund basiert die Bewertung der Sicherheit und Verträglichkeit von Lenvatinib im vorliegenden Dossier allein auf den Analysen der Zeit bis zum Auftreten des ersten Ereignisses.

Die Ergebnisse zur Analyse der Sicherheit und Verträglichkeit ergaben für die Kategorien SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 weder einen geringeren noch größeren Schaden für Lenvatinib gegenüber Sorafenib. Bezüglich der Kategorie UE zeigte sich bei der Analyse eine signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,84; 95%-KI [0,74; 0,95]).

Ergänzend wurden die Ergebnisse der Analysen der UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, dargestellt. Alle UE mit CTCAE Grad 3 oder 4, mit Ausnahme der UE hepatische Enzephalopathie und palmar-plantares Erythrodysästhesiesyndrom, ergaben weder einen geringeren noch größeren Schaden für Lenvatinib gegenüber Sorafenib. In Bezug auf die Zeit bis zum ersten Auftreten einer hepatischen Enzephalopathie mit CTCAE Grad 3 oder 4 ergab sich ein statistisch signifikanter Nachteil für Lenvatinib gegenüber Sorafenib (HR=3,01; 95 5-KI [1,29; 7,02]). Bezüglich der Zeit bis zum ersten Auftreten eines palmar-plantaren Erythrodysästhesiesyndroms mit CTCAE Grad 3 oder 4 ergab sich bei der Analyse eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,23; 95 5-KI [0,13; 0,42]). Alle betrachteten Einzel-UE sind im Rahmen der Behandlung des HCC vorhersehbar und beherrschbar. Die medizinischen

Maßnahmen, die beim Auftreten der jeweiligen Einzel-UE angezeigt sind, werden in der Fachinformation ausführlich beschrieben (Eisai 2018a).

Zusammenfassend zeigte Lenvatinib gegenüber Sorafenib hinsichtlich der Kategorien SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 weder einen geringeren noch größeren Schaden. Bezüglich der Kategorie UE zeigte sich eine signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib. Bei den ergänzenden Analysen der UE mit CTCAE Grad 3 oder 4, deren Inzidenzrate bei ≥ 2% in mindestens einem Behandlungsarm lag, ergab sich insgesamt weder ein geringerer noch ein größerer Schaden für Lenvatinib gegenüber Sorafenib.

Subgruppenanalyse

Bei den rein explorativen Subgruppenanalysen ist zu berücksichtigen, dass ein erhöhtes Auftreten von falsch positiven Ergebnissen aufgrund der Vielzahl der Analysen und des nicht für multiples Testen adjustierten Signifikanzniveaus zu erwarten ist.

Für den Endpunkt OS wurde kein Beleg für eine Interaktion festgestellt. Bezüglich PFS zeigte sich ein konsistentes Bild über alle Subgruppen hinweg, und somit bestätigten die Subgruppenanalysen die Überlegenheit von Lenvatinib gegenüber Sorafenib. Für die Fragebögen EORTC QLQ-C30 und -HCC18 (Morbidität) zeigte sich ebenfalls ein konsistentes Bild über alle Subgruppen. Die gefundenen Interaktionen sind ausschließlich quantitativ mit gleichgerichteten Effekten oder für alle Subgruppen nicht signifikant, so dass ein zufälliges Signal nicht ausgeschlossen werden kann. Aus diesem Grund werden zur Ableitung des Zusatznutzens hinsichtlich der Wirksamkeitsendpunkte OS und PFS und der Fragebögen EORTC QLQ-C30 und -HCC18 die Ergebnisse der Gesamtpopulation herangezogen.

Auch für die Fragebögen EORTC QLQ-C30 und -HCC18 (gesundheitsbezogene Lebensqualität) zeigte sich ein konsistentes Bild über alle Subgruppen. Die gefundenen Interaktionen sind ebenfalls ausschließlich quantitativ mit gleichgerichteten Effekten oder für alle Subgruppen nicht signifikant, so dass von einem zufälligen Signal ausgegangen werden kann. Zur Ableitung des Zusatznutzens hinsichtlich der Endpunkte zur gesundheitsbezogenen Lebensqualität EORTC QLQ-C30 und -HCC18 (gesundheitsbezogene Lebensqualität) werden deshalb die Ergebnisse der Gesamtpopulation herangezogen.

Für die Endpunkte zu Sicherheit und Verträglichkeit finden sich ebenfalls vereinzelt Belege für Interaktionen zwischen der Behandlung und den subgruppenbildenden Merkmalen. Auch hier sind die gefundenen Interaktionen ausschließlich quantitativ mit gleichgerichteten Effekten oder für alle Subgruppen nicht signifikant, so dass von einem zufälligen Signal ausgegangen werden kann. Zur Ableitung des Zusatznutzens hinsichtlich der Sicherheit und Verträglichkeit werden aus diesem Grund die Ergebnisse der Gesamtpopulation herangezogen.

Zusammenfassend konnte hinsichtlich OS, der Morbiditätsendpunkte, der Endpunkte zur gesundheitsbezogenen Lebensqualität und der Endpunkte zu Sicherheit und Verträglichkeit keine hinreichende Evidenz für eine Effektmodifikation in der Studie 304 identifiziert werden.

Daher basiert die Ableitung des Zusatznutzens von Lenvatinib im Vergleich zu Sorafenib für alle berücksichtigten Endpunkte auf den Ergebnissen der Gesamtpopulation.

4.3.3 Weitere Unterlagen

4.3.3.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.3.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Es wurden keine indirekten Vergleiche durchgeführt.

4.3.3.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Nicht zutreffend.

4.3.3.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Nicht zutreffend.

Tabelle 4-71: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	**<Mortalität>**	**<Gesundheits-bezogene Lebensqualität>**	**<Endpunkt>**	**<Endpunkt>**	**<Endpunkt>**
<Studie 1>	nein	ja	ja	ja	nein

4.3.3.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Nicht zutreffend.

Tabelle 4-72: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	**Studie**	**Intervention**	**<Vergleichs-therapie 1>**	**<Vergleichs-therapie 2>**	**<Vergleichs-therapie 3>**
1	<Studie 1>	●		●	●
2	<Studie 2>	●		●	
	<Studie 3>	●		●	
1	<Studie 4>		●	●	●
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Nicht zutreffend.

Tabelle 4-73: Operationalisierung von <Endpunkt xxx>

Studie	**Operationalisierung**
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Nicht zutreffend.

Tabelle 4-74: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Nicht zutreffend.

Tabelle 4-75: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	**Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)**
<Studie 1>	

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.3.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.3.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.3.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.3.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-76: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.3.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.3.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-77: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-78: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.3.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.3.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.3.1 und 4.3.3.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.3.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.3.1 und 4.3.3.2 aufgeführt sind. ***Strukturieren Sie diesen***

Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). *Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.3.3.2 harakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

4.3.3.3.3 Ergebnisse aus weiteren Untersuchungen

4.3.3.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-79: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.3.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.3.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.3.1, 4.3.3.2 und 4.3.3.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 0 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Nachfolgend wird die Aussagekraft der Nachweise des Zusatznutzens von Lenvatinib gegenüber Sorafenib dargelegt.

Die Studie 304 entspricht nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie der Evidenzstufe Ib (G-BA 2018b). Das Verzerrungspotenzial auf Studienebene ist insgesamt als niedrig einzustufen. Die Randomisierung wurde adäquat durchgeführt und die Gruppenzuteilung rein zufällig mittels eines zentralen unabhängigen interaktiven Voice and Web Response Systems (IxRS) realisiert. Die statistischen Methoden entsprechen den Methoden einer Überlebenszeitanalyse und werden mit den entsprechenden Effektschätzern dargestellt. Die Subgruppenanalysen wurden a priori als auch post-hoc

definiert und gemäß den festgelegten Termini ausgewertet und umfassend dargestellt. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP präspezifiziert waren oder – wie bei den Endpunkten zu Sicherheit und Verträglichkeit und den Subgruppenanalysen – gemäß Verfahrensordnung des G-BA berichtet werden (z.B. die Berechnung der Effektmaße für den Vergleich der Behandlungsgruppen).

Das Verzerrungspotenzial für den Endpunkt OS ist als niedrig einzustufen, da die Beeinflussung hinsichtlich des Ergebnisses auch bei einem Open-Label-Studiendesign durch das Wissen um die Behandlung weder durch den Patienten noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Zudem handelt es sich beim OS um einen patientenrelevanten Endpunkt, welcher ein herausragendes Therapieziel darstellt und als gesicherter patientenrelevanter Endpunkt gilt (IQWiG 2011).

Für den Endpunkt PFS zeigt sich ebenfalls ein niedriges Verzerrungspotenzial, da der Endpunkt mit klar definierten, objektiven Kriterien erfasst wurde (mRECIST). Da die Bewertung der Progression in der Hauptanalyse durch ein verblindetes unabhängiges Reviewboard erfolgte, ist keine Beeinflussung aufgrund des Open-Label-Studiendesigns zu erwarten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor.

Das Verzerrungspotential für den Endpunkt EORTC QLQ-C30 ist als hoch einzustufen, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95%. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor.

Das Verzerrungspotential für den Endpunkt EORTC QLQ-HCC18 ist ebenfalls als hoch einzustufen, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95%. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor.

Auch das Verzerrungspotential für den Endpunkt EQ-5D ist als hoch einzustufen, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp

unter dem Zielwert von 95%. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor.

Für die Endpunkte UE, SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 zeigte sich ein niedriges Verzerrungspotenzial, da eine Beeinflussung hinsichtlich des Ergebnisses der Endpunkte auch bei einem Open-Label-Studiendesign durch das Wissen um die Behandlung weder durch den Patienten noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis des SAS, in das alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhalten haben. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor.

Für alle Effektschätzer sind die zugehörigen KI dargestellt, auch für statistisch nicht signifikante Ergebnisse. Die Ergebnisdarstellung gibt keinen Anlass zur Vermutung, dass Ergebnisse selektiv berichtet werden. Es kann daher von einer hohen Ergebnissicherheit ausgegangen werden.

Anhand einer bibliografischen Literaturrecherche (Stand 13. Juli 2018), einer Recherche in Studienregistern (Stand 06. Juli 2018) sowie einer Sichtung der Studien des pharmazeutischen Unternehmers (Stand 01. August 2018) wurde gewährleistet, dass jedwede adäquate Evidenz in die Beurteilung eines Zusatznutzens einfließt und sichergestellt ist, dass keine Daten aus abgebrochenen Studien vorliegen, die einen Einfluss auf die Evidenz haben könnten.

Die Kriterien, welche in der EMA-Guidance „Points to consider on application with 1. Meta-analyses; 2. One pivotal study“ festgelegt sind, können, durch die Erteilung einer Zulassung auf Basis einer Phase-III-Studie als gegeben betrachtet werden. Damit erfüllt die Studie 304 die Anforderungen der EMA vollumfänglich und rechtfertigt eine Berücksichtigung in der Kategorie „hohe Ergebnissicherheit“. Durch die Verfügbarkeit einer RCT mit Evidenzgrad 1b für den Nachweis eines Zusatznutzens sind somit die Kriterien für die Ergebnissicherheit „Hinweis“ erfüllt.

Für die Aussagen zum Zusatznutzen von Lenvatinib als Monotherapie für die Behandlung von erwachsenen Patienten mit fortgeschrittenem oder inoperablem HCC, die zuvor noch keine systemische Therapie erhalten haben gegenüber der zweckmäßigen Vergleichstherapie Sorafenib wird basierend auf den im Dossier vorgelegten Daten bezüglich der Aussagekraft und Ergebnissicherheit der Status eines Hinweises beansprucht.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 0 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Zusatznutzenableitung für Lenvatinib gegenüber der ZVT Sorafenib

Im Folgenden wird der medizinische Zusatznutzen von Lenvatinib, einschließlich dessen Wahrscheinlichkeit und Ausmaß, bewertet. Die Bestimmung des Ausmaßes des Zusatznutzens beruht auf den gemäß der AM-NutzenV festgelegten Kategorien (BMG 2014). Dabei wird für jeden Endpunkt das Ausmaß des Zusatznutzens anhand der Kriterien der AM-NutzenV unter Berücksichtigung des Schweregrads des entsprechenden Ereignisses hergeleitet. Abschließend erfolgt die Bestimmung des Ausmaßes des Zusatznutzens von Lenvatinib im Vergleich zu Sorafenib bei der Betrachtung der Zusammenfassung.

Der Zusatznutzen von Lenvatinib im Vergleich zu Sorafenib wird auf Basis der multizentrischen, randomisierten, Open-Label durchgeführten Nicht-Unterlegenheits- Phase-III-Studie 304 (Evidenzstufe 1b) bewertet. Die Ableitung des Ausmaßes des Zusatznutzens basiert auf der Betrachtung des Zusatznutzens von Lenvatinib gegenüber der zweckmäßigen Vergleichstherapie Sorafenib hinsichtlich der patientenrelevanten Wirksamkeitsendpunkte OS,

PFS und der Fragebögen EORTC QLQ-C30, EORTC QLQ-HCC18 und EQ-5D VAS sowie hinsichtlich patientenrelevanter Endpunkte zu Sicherheit und Verträglichkeit.

Zunächst wird dabei der Zusatznutzen in der Kategorie Mortalität anhand des numerischen Vorteils von Lenvatinib gegenüber Sorafenib in der Hauptanalyse und anhand der statistisch signifikanten Überlegenheit von Lenvatinib gegenüber Sorafenib in den adjustierten Analysen des Endpunkts OS analysiert. Daran anschließend wird der Zusatznutzen anhand der statistischen Überlegenheit von Lenvatinib gegenüber Sorafenib in der Kategorie Morbidität hinsichtlich des Endpunkts PFS und der Skalen der Fragebögen EORTC QLQ-C30 und EORTC QLQ-HCC18 und des Fragebogens EQ-5D VAS betrachtet. Wie auch schon in Abschnitt 4.2.5.2 diskutiert, handelt es sich bei PFS um einen Endpunkt, welcher hinsichtlich der vorliegenden Indikation als unmittelbar und direkt patientenrelevant zu betrachten ist. Begründet ist dies darin, dass eine progressionsfreie Zeit der in den meisten Fällen tödlich verlaufenden Erkrankung einen direkt erlebbaren Erfolg für den Patienten darstellt und einhergehende Ängste zumindest zeitweise lindert. In der Kategorie gesundheitsbezogene Lebensqualität wird der Zusatznutzen anhand der statistischen Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Skalen zur gesundheitsbezogenen Lebensqualität des Fragebogens EORTC QLQ-C30 und der Skalen zur gesundheitsbezogenen Lebensqualität des Fragebogens EORTC QLQ-HCC18 abgeleitet. Bei der Beurteilung des Zusatznutzens von Lenvatinib gegenüber Sorafenib hinsichtlich Sicherheit und Verträglichkeit werden für die Endpunkte Time-To-First-Event-Analysen herangezogen.

Mortalität

Für den Wirksamkeitsendpunkt OS zeigte sich bezüglich des HR ein numerischer Vorteil für Lenvatinib gegenüber Sorafenib basierend auf dem stratifizierten Cox-Proportional-Hazard-Modell (HR=0,92; 95%-KI [0,79; 1,06]; p=0,2526). Da die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag, konnte hiermit die Nicht-Unterlegenheit von Lenvatinib gegenüber Sorafenib belegt werden, was einen Überlebensvorteil von Lenvatinib gegenüber Placebo impliziert. Die mediane Überlebensdauer unter Lenvatinib betrug 13,6 Monate (95%-KI [12,1; 14,9]), unter Sorafenib betrug sie 12,3 Monate (95%-KI [10,4; 13,9]). Damit ergibt sich ein absoluter Vorteil im medianen Überleben von 1,3 Monaten zugunsten von Lenvatinib. In Bezug auf die AFP adjustierte und multivariate adjustierte Analyse zeigte Lenvatinib im Vergleich zu Sorafenib sogar eine statistisch signifikante Überlegenheit im OS. Die in den adjustierten Analysen nachgewiesene statistische Signfikanz kann auf die Imbalance in den Behandlungsarmen bezüglich des für das HCC prognostischen Faktors AFP zurückgeführt werden. Unter Berücksichtigung dieser Ergebnisse zeigte sich somit gemäß § 5 Absatz 7 der AM-NutzenV ein nicht-quantifizierbarer Zusatznutzen für den patientenrelevanten Endpunkt OS.

Morbidität

In der Hauptanalyse des Endpunkts PFS auf Basis des IIR per mRECIST der Studie 304 zeigte sich ein absoluter Vorteil von 3,7 Monaten im medianen PFS zugunsten von Lenvatinib (7,3 Monate; 95%-KI [5,6; 7,5]) gegenüber Sorafenib (3,6 Monate; 95%-KI [3,6; 3,7]), was einer Verdopplung der progressionsfreien Zeit entspricht. Basierend auf dem Cox-Proportional-

Hazard-Modell zeigte Lenvatinib im Vergleich zu Sorafenib eine statistisch signifikante und klinisch relevante Reduktion des Risikos für eine Progression (HR=0,64; 95%-KI [0,55; 0,75]; p<0,0001). Die Robustheit des Ergebnisses der Hauptanalyse wird durch die zwei durchgeführten Sensitivitätsanalysen und die ergänzende Analyse zur ORR, die ein mehrfach höheres Ansprechen unter Lenvatinib im Vergleich zu Sorafenib zeigt, bestätigt. Insgesamt zeigte sich bezüglich des PFS eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens und damit gemäß den Kriterien in § 5 Absatz 7 AM-NutzenV ein Hinweis auf einen beträchtlichen Zusatznutzen.

Die Analyse der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur Morbidität des Fragebogens EORTC QLQ-C30 zeigte für die den Patienten beeinträchtigenden und belastenden Symptome Schmerz (HR=0,81; 95%-KI [0,690; 0,949]; p=0,0091) und Diarrhö (HR=0,52; 95%-KI [0,440; 0,623]; p<0,0001) eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib. Die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf der Skala Schmerz betrug dabei 1,9 Monate (95%-KI [1,87; 2,76]) für Lenvatinib im Vergleich zu 1,8 Monate (95%-KI [1,12; 1,87]) für Sorafenib. Die Kaplan-Meier-Kurve für das Symptom Schmerz zeigt eine Trennung der Kurven und einen Vorteil für Lenvatinib über den aufgrund genügend vorliegender Datenpunkte interpretierbaren Beobachtungszeitraum, wobei die Kurven beim Median eng beieinanderliegen. Diese Überlegenheit von Lenvatinib gegenüber Sorafenib ist relevant, da insbesondere Schmerzen für den Patienten direkt spürbar sind und somit eine Vermeidung oder Verzögerung des Auftretens von Schmerzen zum Wohlbefinden des Patienten beiträgt. Bezüglich der Skala Diarrhö betrug der absolute Vorteil 1,9 Monate in der medianen Zeit bis zur ersten Verschlechterung um 10 Punkte zugunsten von Lenvatinib (4,6 Monate; 95%-KI [3,71; 5,52]) im Vergleich zu Sorafenib (2,7 Monate; 95%-KI [1,87; 2,79]). Diarrhö wird häufig als äußerst unangenehm empfunden und schränkt das tägliche Leben des Patienten erheblich ein, so dass auch hier eine Vermeidung oder Verzögerung des Auftretens von Diarrhö unmittelbar zum Wohlbefinden des Patienten beiträgt. Hinsichtlich der Symptome Fatigue, Übelkeit und Erbrechen, Dyspnoe, Schlaflosigkeit, Appetitminderung und Obstipation zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib. Insgesamt zeigte sich somit bezüglich des Symptoms Schmerz eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapierelevanten Nutzens und damit gemäß den Kriterien in § 5 Absatz 7 der AM-NutzenV ein Hinweis auf einen geringen Zusatznutzen. Bezüglich des Symptoms Diarrhö zeigte sich eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte große Verbesserung des therapierelevanten Nutzens und damit gemäß den Kriterien in § 5 Absatz 7 der AM-NutzenV ein Hinweis auf einen erheblichen Zusatznutzen.

Auch im Hinblick auf die Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den Skalen zur Morbidität des Fragebogens EORTC QLQ-HCC18 konnten die Patienten unter der Behandlung mit Lenvatinib für alle Symptome vergleichbare Therapieerfolge erzielen wie unter der Behandlung mit Sorafenib. Insgesamt zeigte sich gemäß den Kriterien in § 5 Absatz 7 der AM-NutzenV kein Hinweis auf einen Zusatznutzen für die Skalen zur Morbidität des EORTC QLQ-HCC18.

Die Analyse des Fragebogens EQ-5D bezüglich der Zeit bis zur ersten Verschlechterung um mindestens 7 oder 10 mm auf der VAS zeigte für die Patienten unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter Sorafenib. Insgesamt zeigte sich gemäß den Kriterien in § 5 Absatz 7 der AM-NutzenV kein Hinweis auf einen Zusatznutzen für die EQ-5D VAS.

Gesundheitsbezogene Lebensqualität

Bei der Analyse des Rollenfunktionsniveaus des Fragebogens EORTC QLQ-C30 zeigte sich bezüglich der ersten Verschlechterung um mindestens 10 Punkte auf der Skala eine statistisch signifikante und patientenrelevante Überlegenheit von Lenvatinib gegenüber Sorafenib (HR=0,82; 95 5-KI [0,696; 0,962]; p=0,0148). Die mediane Zeit betrug dabei 2,0 Monate (95%-KI [1,87; 2,83]) für Lenvatinib im Vergleich zu 1,9 Monate (95%-KI [1,84; 1,97]) für Sorafenib. Die Kaplan-Meier-Kurve für das Rollenfunktionsniveau zeigt eine Trennung der Kurven und einen Vorteil für Lenvatinib über den aufgrund genügend vorliegender Datenpunkte interpretierbaren Beobachtungszeitraum, wobei die Kurven beim Median eng beieinanderliegen. Diese Überlegenheit von Lenvatinib gegenüber Sorafenib ist relevant, da insbesondere die Bewältigung der Arbeit, des Alltags und der Freizeitaktivitäten für den Patienten direkt wahrnehmbar ist und somit zum Wohlbefinden des Patienten beiträgt. Hinsichtlich der physischen, emotionalen, kognitiven, sozialen und finanziellen Funktion und des globalen Gesundheitsstatus zeigten sich unter der Behandlung mit Lenvatinib vergleichbare Therapieerfolge wie unter der Behandlung mit Sorafenib. Insgesamt zeigte sich bezüglich des Rollenfunktionsniveaus somit eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapierelevanten Nutzens und damit gemäß den Kriterien in § 5 Absatz 7 der AM-NutzenV ein Hinweis auf einen geringen Zusatznutzen.

Auch die Analysen des Fragenbogens EORTC QLQ-HCC18 bezüglich Körperbild (HR=0,79; 95%-KI [0,671; 0,932]; p=0,0051) und Ernährung (HR=0,80; 95%-KI [0,674; 0,949]; p=0,0102) zeigten eine statistisch signifikante Überlegenheit von Lenvatinib gegenüber Sorafenib hinsichtlich der Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte auf den entsprechenden Skalen. Bezüglich Körperbild betrug die mediane Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte 2,8 Monate (95%-KI [2,73; 3,68]) für Lenvatinib im Vergleich zu 1,9 Monate (95%-KI [1,87; 2,73]) für Sorafenib. Insbesondere die Vermeidung oder Verzögerung des Auftretens einer Verschlechterung des Körperbildes, wie der Verlust von Muskelmasse an Armen und Beinen und das Anschwellen des Bauches, führt unmittelbar zu einer Verbesserung der Lebensqualität. Bezüglich des Items Ernährung betrug die Zeit 4,3 Monate (95%-KI [3,68; 5,52]) für Lenvatinib im Vergleich zu 2,8 Monate (95%-KI [2,07; 3,02]) für Sorafenib. Diese Überlegenheit von Lenvatinib gegenüber Sorafenib ist relevant, da insbesondere die Vermeidung oder Verzögerung eines fehlenden Geschmacksinns, eines schnellen Sättigungsgefühls, einer unzureichenden Nahrungsaufnahme und ein damit einhergehendes niedriges Körpergewicht für den Patienten direkt wahrnehmbar ist und somit zum Wohlbefinden des Patienten beiträgt. Hinsichtlich des Sexuallebens zeigten sich für Lenvatinib vergleichbare Ergebnisse wie für Sorafenib. Insgesamt zeigte sich für Körperbild und Ernährung eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte

moderate und nicht nur geringfügige Verbesserung des therapierelevanten Nutzens und damit gemäß den Kriterien in § 5 Absatz 7 der AM-NutzenV ein Hinweis auf einen geringen Zusatznutzen.

Sicherheit und Verträglichkeit

Die Ergebnisse zur Analyse der Endpunkte zu Sicherheit und Verträglichkeit ergaben für die Kategorien SUE, UE, die zum Therapieabbruch führten und UE mit CTCAE Grad ≥ 3 weder einen geringeren noch größeren Schaden für Lenvatinib gegenüber Sorafenib. Bezüglich der Kategorie UE zeigte sich bei der Analyse eine Überlegenheit von Lenvatinib gegenüber Sorafenib.

Die nachfolgende Tabelle zeigt die Ableitung der Aussagesicherheit und des Ausmaßes des Zusatznutzens gemäß der AM-NutzenV festgelegten Kategorien.

Tabelle 4-80: Ausmaß des Zusatznutzens von Lenvatinib vs. Sorafenib auf Endpunktebene

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
Mortalität		
OS (Hauptanalyse)	HR=0,92 [0,79; 1,06]; p=0,2526 13,6 vs. 12,3 Monate	**Hinweis / Zusatznutzen mit nicht-quantifizierbarem Ausmaß**
OS (multivariate adjustierte Analyse)	HR=0,82 [0,70; 0,96]	
OS (AFP adjustierte Analyse)	HR=0,85 [0,73; 0,99]	
Morbidität		
PFS (IIR per mRECIST)	HR=0,64 [0,55; 0,75]; p<0,0001 7,3 vs. 3,6 Monate	**Hinweis / Zusatznutzen mit Ausmaß beträchtlich**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Fatigue	HR=0,93 [0,797; 1,085]; p=0,3592 1,8 vs. 1,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Übelkeit und Erbrechen	HR=1,04 [0,857; 1,263]; p=0,6894 6,4 vs. 7,3 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Schmerz	HR=0,81 [0,690; 0,949]; p=0,0091 1,9 vs. 1,8 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Dyspnoe	HR=0,97 [0,797; 1,170]; p=0,7217 6,4 vs. 5,6 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Schlaflosigkeit	HR=1,16 [0,958; 1,398]; p=0,1293 4,6 vs. 6,1 Monate	Kein Hinweis auf einen Zusatznutzen

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert** **mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten** **Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Appetitminderung	HR=1,02 [0,858; 1,201]; p=0,8575 2,8 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Obstipation	HR=1,06 [0,868; 1,301]; p=0,5537 7,3 vs. 9,2 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Diarrhö	HR=0,52 [0,440; 0,623]; p<0,0001 4,6 vs. 2,7 Monate	**Hinweis / Zusatznutzen mit Ausmaß erheblich**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Fatigue	HR=0,94 [0,802; 1,097]; p=0,4216 1,8 vs. 1,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Ikterus	HR=0,95 [0,790; 1,144]; p=0,5934 4,6 vs. 3,7 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Schmerz	HR=1,14 [0,961; 1,349]; p=0,1315 2,8 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Fieber	HR=0,90 [0,752; 1,083]; p=0,2705 5,5 vs. 3,7 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Abdominale Schwellung	HR=0,98 [0,805; 1,197]; p=0,8549 7,4 vs. 7,4 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 7 mm im EQ-5D VAS	HR=0,90 [0,765; 1,056]; p=0,1951 2,8 vs. 1,9 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 mm im EQ-5D VAS	HR=0,89 [0,758; 1,050]; p=0,1713 2,8 vs. 1,9 Monate	Kein Hinweis auf einen Zusatznutzen
Gesundheitsbezogene Lebensqualität		
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Physische Funktion	HR=0,89 [0,756; 1,056]; p=0,1870 3,7 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Rollenfunktionsniveau	HR=0,82 [0,696; 0,962]; p=0,0148 2,0 vs. 1,9 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Emotionale Funktion	HR=1,01 [0,827; 1,235]; p=0,9142 7,4 vs. 6,6 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Kognitive Funktion	HR=1,09 [0,920; 1,291]; p=0,3167 2,8 vs. 3,0 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Soziale Funktion	HR=1,06 [0,899; 1,260]; p=0,4658 2,8 vs. 2,8 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Finanzielle Schwierigkeiten	HR=0,91 [0,735; 1,130]; p=0,3980 12,3 vs. 8,5 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-C30 Globaler Gesundheitsstatus	HR=0,99 [0,841; 1,171]; p=0,9247 2,7 vs. 2,1 Monate	Kein Hinweis auf einen Zusatznutzen
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Körperbild	HR=0,79 [0,671; 0,932]; p=0,0051 2,8 vs. 1,9 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Ernährung	HR=0,80 [0,674; 0,949]; p=0,0102 4,3 vs. 2,8 Monate	**Hinweis / Zusatznutzen mit Ausmaß gering**
Zeit bis zur ersten Verschlechterung um mindestens 10 Punkte im EORTC QLQ-HCC18 Sexualleben	HR=0,94 [0,771; 1,154]; p=0,5700 7,4 vs. 6,7 Monate	Kein Hinweis auf einen Zusatznutzen
Sicherheit und Verträglichkeit		
Zeit bis zum ersten UE	HR=0,84 [0,74; 0,95] 8 vs. 8 Tage	Kein geringerer oder größerer Schaden
Zeit bis zum ersten SUE	HR=1,24 [0,999; 1,53)] 412 vs. 708 Tage	Kein geringerer oder größerer Schaden
Zeit bis zum ersten UE, das zum Therapieabbruch führte	HR=1,10 [0,80; 1,50] NE vs. NE Tage	Kein geringerer oder größerer Schaden
Zeit bis zum ersten UE mit CTCAE Grad ≥ 3	HR=0,96 [0,83; 1,12] 82 vs. 64 Tage	Kein geringerer oder größerer Schaden

Endpunktkategorie Endpunkt	**Effektschätzer [95%-KI]; p-Wert mediane Zeit bis zum Ereignis bzw. Anzahl der Patienten Lenvatinib vs. Sorafenib**	**Ergebnissicherheit / Ausmaß des Zusatznutzens**
AFP: Alpha-Fetoprotein; CTCAE: Common Terminology Criteria for Adverse Events (Allgemeine Terminologiekriterien für unerwünschte Ereignisse); EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; EQ-5D VAS: EuroQoL Five Dimension Health Questionnaire - Visual Analogue Scale; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); HR: Hazard Ratio; IIR: Independent Imaging Review (unabhängige radiologische Bewertung); KI: Konfidenzintervall; mRECIST: modified Response Evaluation Criteria in Solid Tumors; NE: not estimable (nicht berechenbar); OS: Overall Survival (Gesamtüberleben); PFS: Progression Free Survival (progressionsfreies Überleben); SUE: Schwerwiegendes unerwünschtes Ereignis / schwerwiegende unerwünschte Ereignisse; UE: Unerwünschtes Ereignis / unerwünschte Ereignisse		

In Folgenden wird die Gesamtaussage zum Zusatznutzen von Lenvatinib gegenüber Sorafenib tabellarisch dargestellt (Tabelle 4-81).

Tabelle 4-81: Übersicht zur Gesamtaussage zum Zusatznutzen von Lenvatinib vs. Sorafenib

Zusatznutzen	**Größerer Schaden**
erheblich • EORTC QLQ-C30 Symptome (Hinweis) • Diarrhö	erheblich –
beträchtlich • PFS (Hinweis)	beträchtlich –

gering
• EORTC QLQ-C30 Symptome (Hinweis)
  • Schmerz
• EORTC QLQ-C30 QoL (Hinweis)
  • Rollenfunktionsniveau
• EORTC QLQ-HCC18 QoL (Hinweis)
  • Körperbild
  • Ernährung
	gering
	–

nicht-quantifizierbar • OS	nicht-quantifizierbar –
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire; EORTC QLQ-HCC18: European Organization for Research and Treatment of Cancer - HCC-specific Quality of Life Questionnaire; HCC: Hepatocellular Carcinoma (Leberzellkarzinom); OS: Overall Survival (Gesamtüberleben); PFS: Progression-free Survival (progressionsfreies Überleben); QoL: Quality of Life	

Zusammenfassung

Unter Berücksichtigung der für die direkt und unmittelbar patientenrelevanten Wirksamkeitsendpunkte OS und PFS, für die Skalen zur Morbidität des EORTC QLQ-C30 und für die Skalen zur gesundheitsbezogenen Lebensqualität des EORTC QLQ-30 und QLQ-HCC18 gezeigten Verbesserungen des therapierelevanten Nutzens für diese seltene, sehr schwere und praktisch immer unheilbare Tumorerkrankung, für die ein dringender Therapiebedarf bei gleichzeitig sehr stark beschränkter Anzahl an Therapieoptionen besteht, ergibt sich in der Zusammenschau mit dem insgesamt vergleichbaren, vorhersehbaren und beherrschbaren Sicherheits- und Verträglichkeitsprofil für Lenvatinib gegenüber der ZVT Sorafenib ein **Hinweis auf einen erheblichen Zusatznutzen**.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 0 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-82: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit fortgeschrittenem oder inoperablem hepatozellulärem Karzinom (HCC), die zuvor noch keine systemische Therapie erhalten haben.	Hinweis auf einen erheblichen Zusatznutzen

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.3.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.3.2) oder weitere Untersuchungen (Abschnitt 4.3.3.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005[13], Molenberghs 2010[14]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006[15]) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf

[13] *Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.*
[14] *Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.*
[15] *Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.*

den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006[16]) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

– Patientenpopulation
– Intervention
– Kontrolle
– Datenherkunft
– verwendete Methodik
– entsprechende Ergebnisse (abhängig von der Methode)
– Untersuchungen zur Robustheit
– ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht zutreffend.

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

[16] *Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.*

Tabelle 4-83: Liste der eingeschlossenen Studien

Studie	verfügbare Quellen[a]		
	Studienbericht (ja/nein [Zitat])	Registereintrag[c] (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
E7080-G000-304	ja (Eisai Inc. 2017b)	Clinicaltrials.gov NCT00946153 (ClinicalTrialsGOV 2018) EU CTR 2012-002992-33 (EU-Clinical Trials Register 2012) PharmNet.Bund 2012-002992-33 (PharmnetBund 2012) WHO ICTRP NCT01761266 (WHO ICTRP 2018)	Ja (Kudo et al. 2018)

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Aaronson N. K., Ahmedzai S., Bergman B. et al. 1993. *The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology.* Journal of the National Cancer Institute 85 (5), S. 365–376.
2. Abou-Alfa G. K., Cheng A.-L., Meyer T. et al. 2015. *Phase 3 randomized, double-blind, controlled study of cabozantinib (XL184) versus placebo in subjects with hepatocellular carcinoma who have received prior sorafenib (CELESTIAL; NCT01908426).* JCO 33 (3_suppl), S. TPS496-TPS496.
3. Arbeitsgemeinschaft Dermatologische Onkologie (ADO) 2012. *Stellungnahme zur Nutzenbewertung des IQWiG gemäß §35a SGB V*: *"Vemurafenib"*. Verfügbar unter: https://hautkrebs.files.wordpress.com/2012/07/stellungnahme_ado_iqwig_vemurafenib_final.pdf, abgerufen am: 08.12.2017.
4. Bayer Pharma AG 2018. *Fachinformation Nexavar® 200 mg Filmtabletten; Wirkstoff: Sorafenib*. *Stand Juni 2018*.

5. Berg P., Book K., Dinkel A. et al. 2011. *Progredienzangst bei chronischen Erkrankungen.* Psychother Psych Med 61, S. 32–37.
6. Bolondi L., Burroughs A., Dufour J.-F. et al. 2012. *Heterogeneity of patients with intermediate (BCLC B) Hepatocellular Carcinoma: proposal for a subclassification to facilitate treatment decisions.* Seminars in Liver Disease 32 (4), S. 348–359.
7. Bruix J., Qin S., Merle P. et al. 2017. *Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial.* The Lancet 389 (10064), S. 56–66.
8. Bruix J., Sherman M., Llovet J. M. et al. 2001. *Clinical Management of Hepatocellular Carcinoma. Conclusions of the Barcelona-2000 EASL Conference.* Journal of Hepatology 35 (3), S. 421–430.
9. Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) 2011. *Regulators' Assessment for Marketing Authorization Decisions*. Verfügbar unter: https://www.g-ba.de/downloads/17-98-3002/BfArM.pdf, abgerufen am: 30.07.2018.
10. Bundesministerium für Gesundheit (BMG) 2014. *Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - Am-NutzenV),*
11. Chie W.-C., Blazeby J. M., Hsiao C.-F. et al. 2012. *International cross-cultural field validation of an European Organization for Research and Treatment of Cancer questionnaire module for patients with primary liver cancer, the European Organization for Research and Treatment of Cancer quality-of-life questionnaire HCC18.* Hepatology 4 (55), S. 1122–1129.
12. ClinicalTrialsGOV 2018. *A Multicenter, Open-Label, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib (E7080) Versus Sorafenib in First-Line Treatment of Subjects With Unresectable Hepatocellular Carcinoma*: *NCT01761266 / E7080-G000-304*. Verfügbar unter: https://ClinicalTrials.gov/show/NCT01761266, abgerufen am: 31.07.2018.
13. Cocks K., King M. T., Velikova G. et al. 2011. *Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology 29 (1), S. 89–96.
14. Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO) 2013. *Nutzenbewertung von Arzneimitteln der Onkologie und Hämatologie 2013*: *Eine Standortbestimmung ; Teil 1: Postitionspapier, Teil 2: Manual,* 2013. Aufl. DGHO, Berlin, S. 123.
15. Deutsche Gesellschaft für Verdauungs- und Stoffwechselkrankheiten (DGVS) 2013. *Diagnostik und Therapie des hepatozellulären Karzinoms, Langversion 1.0*: *AWMF-Registernummer: 032/053OL*. *Leitlinie*. Verfügbar unter: https://www.leitlinienprogramm-

onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/HCC/S3-HCC-OL-Langversion-V1.0.pdf, abgerufen am: 24.07.2018.

16. Eisai Europe Ltd. (Eisai) 2018a. *Fachinformation Lenvima® 4 mg/10 mg Hartkapseln; Wirkstoff: Lenvatinib*. *Stand August 2018*.
17. Eisai Europe Ltd. (Eisai) 2018b. *Summary of ongoing, completed, or discontinued efficacy and safety clinical studies with lenvatinib in HCC subjects; Status 01 Aug 2018*. *Data on file*.
18. Eisai Inc. 2016. *Statistischer Analyseplan der Studie E7080-G000-304*. *Data on file*.
19. Eisai Inc. 2017a. *Post Hoc Retrospective Independent Imaging Review Report*. *Data on file*.
20. Eisai Inc. 2017b. *Studienbericht der Studie E7080-G000-304*. *Data on file*.
21. Eisai Inc. 2018a. *MVA Report: Covariate Adjusted Analysis of Lenvatinib Versus Sorafenib in Treatment of Unresectable Hepatocellular Carcinoma*. *Data on file*.
22. Eisai Inc. 2018b. *Prior treatment and baseline TACE contraindications in BCLC B patients - Study 304*. *Data on file*.
23. Eisai Inc. 2018c. *Zusatzanalysen zur Studie 304. Data on file*.
24. EU-Clinical Trials Register 2012. *A Multicenter, Randomized, Open-Label, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib (E7080) Versus Sorafenib in First-Line Treatment of Subjects With Unresectable Hepatocellular*: *2012-002992-33 / E7080-G000-304*. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002992-33, abgerufen am: 31.07.2018.
25. European Association for the Study of the Liver (EASL) 2018. *EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma.* Journal of Hepatology 69 (1), S. 182–236.
26. European Medicines Agency (EMA) 2001. *Points to consider on application with 1. meta-analyses; 2. one pivotal study*. Verfügbar unter: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003657.pdf, abgerufen am: 30.07.2018.
27. European Medicines Agency (EMA) 2012. *Guideline on the evaluation of anticancer medicinal products in man*. Verfügbar unter: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137128.pdf, abgerufen am: 30.07.2018.
28. Food and Drug Administration (FDA) 2007. *Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics*. Verfügbar unter: https://www.fda.gov/downloads/Drugs/Guidances/ucm071590.pdf, abgerufen am: 08.12.2017.
29. Gemeinsamer Bundesausschuss (G-BA) 2014. *Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-*

RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Ruxolitinib. Verfügbar unter: https://www.g-ba.de/downloads/40-268-2996/2014-11-06_AM-RL-XII_Ruxolitinib_2014-05-15-D-108_TrG.pdf, abgerufen am: 08.12.2017.

30. Gemeinsamer Bundesausschuss (G-BA) 2018a. *Niederschrift (finale Fassung) zum Beratungsgespräch gemäß §8 Abs. 1 AM-NutzenV Beratungsanforderung 2017-B-221*. *Data on file*.
31. Gemeinsamer Bundesausschuss (G-BA) 2018b. *Verfahrensordnung des Gemeinsamen Bundesausschusses in der Fassung vom 18. Dezember 2008 veröffentlicht im Bundesanzeiger Nr. 84a (Beilage) vom 10. Juni 2009 - in Kraft getreten am 1. April 2009*: *zuletzt geändert am 16. März 2018 - veröffentlicht im Bundesanzeiger BAnz AT 04.07.2018 B1 - in Kraft getreten am 5. Juli 2018.* Verfügbar unter: https://www.g-ba.de/downloads/62-492-1614/VerfO_2018-03-16_iK-2018-07-05.pdf, abgerufen am: 08.08.2018.
32. Giesinger J. M., Kieffer J. M., Fayers P. M. et al. 2016. *Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust.* Journal of Clinical Epidemiology 69, S. 79–88.
33. Herschbach P., Dankert A., Duran-Atzinger G. et al. 2001. *Diagnostik von Progredienzangst*: *Entwicklung eines Fragebogens zur Erfassung von Progredienzangst bei Patienten mit Krebserkrankungen, Diabetes mellitus und entzündlich-rheumatischen Erkrankungen in der Rehabilitation.* Verfügbar unter: http://forschung.deutsche-rentenversicherung.de/ForschPortalWeb/rehaDoc.pdf?rehaid=82EF8B8F899B55CEC1256E6A003B729D, abgerufen am: 15.02.2018.
34. Herschbach P. und Dinkel A. 2014. *Fear of Progression*. Verfügbar unter: https://link.springer.com/chapter/10.1007%2F978-3-642-40187-9_2#citeas, abgerufen am: 15.03.2018.
35. Hesse M., Heydweiller K., Mücke M. et al. 2016. *Angst bei fortschreitenden Erkrankungen.* Dtsch med Wochenschr 141 (17), S. 1229–1234.
36. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2011. *Aussagekraft von Surrogatendpunkten in der Onkologie*: *Rapid Report, Version 1.1*. *IQWiG-Berichte – Jahr: 2011 Nr. 80*. Verfügbar unter: https://www.iqwig.de/download/A10-05_Rapid_Report_Version_1-1_Surrogatendpunkte_in_der_Onkologie.pdf, abgerufen am: 15.09.2017.
37. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2013. *Crizotinib - Nutzenbewertung gemäß § 35a SGB V*. Verfügbar unter: https://www.g-ba.de/downloads/92-975-168/2013-02-13_A12-15_Crizotinib_Nutzenbewertung%20%C2%A7%2035a%20SGB%20V.pdf, abgerufen am: 02.12.2017.
38. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2014. *Afatinib - Nutzenbewertung gemäß § 35a SGB V*. Verfügbar unter: https://www.g-

ba.de/downloads/92-975-403/2014-02-13_A13-41_Afatinib_Nutzenbewertung-35a-SGB-V.pdf, abgerufen am: 08.12.2017.

39. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2017. *Allgemeine Methoden*: *Version 5.0 vom 10.07.2017*. Verfügbar unter: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf, abgerufen am: 15.09.2017.
40. Kudo M. 2018. *Lenvatinib May Drastically Change the Treatment Landscape of Hepatocellular Carcinoma.* Liver cancer 7 (1), S. 1–19.
41. Kudo M., Arizumi T., Ueshima K. et al. 2015. *Subclassification of BCLC B Stage Hepatocellular Carcinoma and Treatment Strategies: Proposal of Modified Bolondi's Subclassification (Kinki Criteria).* Digestive diseases (Basel, Switzerland) 33 (6), S. 751–758.
42. Kudo M., Finn R. S., Qin S. et al. 2018. *Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial.* The Lancet 391 (10126), S. 1163–1173.
43. Lencioni R. und Llovet J. M. 2010. *Modified RECIST (mRECIST) Assessment for Hepatocellular Carcinoma.* Seminars in Liver Disease 30 (1), S. 52–60.
44. Li L., Mo F., Chan S. et al. 2017. *Prognostic values of EORTC QLQ-C30 and QLQ-HCC18 index-scores in patients with hepatocellular carcinoma – clinical application of health-related quality-of-life data.* BMC cancer 17 (8), S. 1–11.
45. Medac 2013. *Fachinformation Adrimedac® 2 mg/ml Infusionslösung; Wirkstoff: Doyorubicinhydrochlorid*. *Stand September 2013*.
46. Mehnert A., Herschbach P., Berg P. et al. 2006. *Progredienzangst bei Brustkrebspatientinnen - Validierung der Kurzform des Progredienzangstfragebogens PA-F-KF*: *Fear of progression in breast cancer patients - Validation of the short form of the Fear of Progression Questionnaire (FoP-Q-SF).* Zeitschrift für Psychosomatische Medizin und Psychotherapie 52 (3), S. 274–288.
47. Moher D., Hopewell S., Schulz K. F. et al. 2010. *CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials.* Journal of Clinical Epidemiology 63 (8), S. e1-e37.
48. Osoba D., Rodrigues G., Myles J. et al. 1998. *Interpreting the significance of changes in health-related quality-of-life scores.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology 16 (1), S. 139–144.
49. PharmnetBund 2012. *A Multicenter, Randomized, Open-Label, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib (E7080) Versus Sorafenib in First-Line Treatment of Subjects With Unresectable Hepatocellular Carcinoma*: *2012-002992-33 / E7080-G000-304*. Verfügbar unter: WWW.PHARMNET-BUND.DE, abgerufen am: 31.07.2018.

50. Pickard A. S., Neary M. P. und Cella D. 2007. *Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer.* Health and Quality of Life Outcomes 5, S. 70.

51. Pickard A. S., Ray S., Ganguli A. et al. 2012. *Comparison of FACT- and EQ-5D–Based Utility Scores in Cancer.* Value in Health 15 (2), S. 305–311.

52. Raoul J. L., Sangro B., Forner A. et al. 2011. *Evolving strategies for the management of intermediate-stage hepatocellular carcinoma: Available evidence and expert opinion on the use of transarterial chemoembolization.* Cancer Treat Rev 2011 (37), S. 212–220.

53. Schulz K. F., Altmann D. G. und Moher D. 2010. *CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.* BMC Medicine 8 (18), S. e1–9.

54. SCO:SSiS 2018. *Zusatzanalysen zur Studie 304*. *Data on file*.

55. Sieghart W., Hucke F. und Peck-Radosavljevic M. 2015. *Transarterial chemoembolization: modalities, indication, and patient selection.* Journal of Hepatology 62 (5), S. 1187–1195.

56. Teckle P., McTaggart-Cowan H., Van der Hoek, K. et al. 2013. *Mapping the FACT-G cancer-specific quality of life instrument to the EQ-5D and SF-6D*. Verfügbar unter: https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-11-203, abgerufen am: 22.01.2018.

57. Thomas M. B., Jaffe D., Choti M. M. et al. 2010. *Hepatocellular Carcinoma: Consensus Recommendations of the National Cancer Institute Clinical Trials Planning Meeting.* JCO 28 (25), S. 3994–4005.

58. Verslype C., Rosmorduc O. und Rougier P. 2012. *Hepatocellular carcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Annals of oncology : official journal of the European Society for Medical Oncology 23 Suppl 7 (23 (Supplement 7)), S. vii41-8.

59. WHO ICTRP 2018. *A Multicenter, Open-Label, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib (E7080) Versus Sorafenib in First-Line Treatment of Subjects With Unresectable Hepatocellular Carcinoma*: *NCT01761266 / E7080-G000-304*. Verfügbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01761266, abgerufen am: 31.07.2018.

60. Wong S. S.L., Wilszynski N. L. und Haynes R. B. 2006. *Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE*.* Journal of the Medical Library Association 94 (4), S. 451–455.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[17]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

[17] Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	13.07.2018	
Zeitsegment	1988 to 2018 Week 28	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 (Wong et al. 2006)	
#	**Suchbegriffe**	**Ergebnis**
Disease terms		
1	exp liver cell carcinoma/	122461
2	exp liver tumor/ or exp liver cancer/	218020
3	((hepatocellular* or hepatic* or liver*) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).af.	316801
4	hcc.af.	72218
5	(hepatoma*).af.	25829
6	(hepatoma* adj3 malignant*).af.	164
7	(hepatocarcinoma*).af.	4809
8	**1 or 2 or 3 or 4 or 5 or 6 or 7**	**349467**
Study design terms		
9	placebo.mp.	369892
10	"random*".tw.	1264805
11	"double-blind*".tw.	166199
12	**9 or 10 or 11**	**1461753**
Intervention terms		
13	Lenvatinib*.af.	1136
14	exp lenvatinib/	1102
15	E7080.af.	129
16	E-7080.af.	165
17	E 7080.af.	165
18	er 203492-00.af.	1
19	er203492-00.af.	0
20	Lenvatinib mesylat*.af.	7
21	Lenvatinib mesilat*.af.	13
22	"4-(3-chloro-4-(N'-cyclopropylureido)phenoxy)-7-methoxyquinoline-6-carboxamide".af.	4
23	"4-(3-chloro-4-((cyclopropylaminocarbonyl)amino)phenoxy)-7-methoxy-6-quinolinecarboxamide".af.	0
24	"Lenvima*".af.	79

25	"Kisplyx*".af.	3
26	**13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25**	**1186**
Final facet		
27	**8 and 12 and 26**	**104**

Datenbankname	Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) Daily Update	
Suchoberfläche	Ovid	
Datum der Suche	13.07.2018	
Zeitsegment	1946 to present	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 (Wong et al. 2006)	
#	**Suchbegriffe**	**Ergebnis**
Disease terms		
1	exp Liver Neoplasms/	150808
2	((hepatocellular* or liver*) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).af.	207055
3	hcc.af.	43700
4	(hepatoma*).af.	27429
5	hepatocarcinoma.af.	3520
6	(hepatoma* adj3 malignant*).af.	245
7	**1 or 2 or 3 or 4 or 5 or 6**	**217998**
Study design terms		
8	randomized.mp.	735650
9	randomised.mp.	88151
10	random.mp.	321874
11	placebo.mp.	193340
12	randomized controlled trial.pt.	463881
13	**8 or 9 or 10 or 11 or 12**	**1080980**
Intervention terms		
14	Lenvatinib*.af.	292
15	E7080.af.	48
16	E-7080.af.	2
17	E 7080.af.	2
18	"Lenvatinib mesylat*".af.	4
19	"Lenvatinib mesilat*".af.	1
20	"4-(3-chloro-4-(N'-cyclopropylureido)phenoxy)-7-methoxyquinoline-6-carboxamide".af.	1

21	"4-(3-chloro-4-((cyclopropylaminocarbonyl)amino)phenoxy)-7-methoxy-6-quinolinecarboxamide".af.	0
22	Lenvima*.af.	11
23	Kisplyx*.af.	0
24	**14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23**	**301**
Final facet		
25	**7 and 13 and 24**	**10**

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials (Clinical Trials)	
Suchoberfläche	Ovid	
Datum der Suche	13.07.2018	
Zeitsegment	June 2018	
#	**Suchbegriffe**	**Ergebnis**
Disease terms		
1	exp Liver Neoplasms/	2504
2	((hepatocellular* or liver*) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).af.	6149
3	hcc.af.	2284
4	(hepatoma*).af.	114
5	hepatocarcinoma.af.	55
6	(hepatoma* adj3 malignant*).af.	4
7	**1 or 2 or 3 or 4 or 5 or 6**	**6422**
Intervention terms		
8	Lenvatinib*.af.	134
9	E7080.af.	35
10	E-7080.af.	0
11	E 7080.af.	0
12	Lenvatinib mesylat*.af.	0
13	lenvatinib mesilate*.af.	1
14	"4-(3-chloro-4-(N'-cyclopropylureido)phenoxy)-7-methoxyquinoline-6-carboxamide".af.	0
15	"4-(3-chloro-4-((cyclopropylaminocarbonyl)amino)phenoxy)-7-methoxy-6-quinolinecarboxamide".af.	0
16	"Lenvima*".af.	2
17	“Kisplyx*”.af.	0
18	**8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17**	**134**
Final facet		
19	**7 and 18**	**21**

Datenbankname	EBM Reviews - Cochrane Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	13.07.2018	
Zeitsegment	2005 to July 11, 2018	
#	**Suchbegriffe**	**Ergebnis**
Disease terms		
1	((hepatocellular* or liver*) adj3 (carcinoma* or adenocarcinoma* or cancer* or neoplasm* or tumo?r*)).af.	311
2	hcc.af.	45
3	(hepatoma*).af.	24
4	hepatocarcinoma.af.	7
5	(hepatoma* adj3 malignant*).af.	2
6	**1 or 2 or 3 or 4 or 5**	**332**
Intervention terms		
7	Lenvatinib*.af.	0
8	E7080.af.	0
9	E-7080.af.	0
10	E 7080.af.	0
11	Lenvatinib mesylat*.af.	0
12	lenvatinib mesilat*.af.	0
13	"4-(3-chloro-4-(N'-cyclopropylureido)phenoxy)-7-methoxyquinoline-6-carboxamide".af.	0
14	"4-(3-chloro-4-((cyclopropylaminocarbonyl)amino)phenoxy)-7-methoxy-6-quinolinecarboxamide".af.	0
15	"Lenvima*".af.	0
16	"Kisplyx*".af.	0
17	**9 or 10 or 11 or 12 or 13 or 14 or 15 or 16**	**0**
Final facet		
18	**6 and 17**	**0**

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. clinicaltrials.gov), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in ClinicalTrials.gov

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	06.07.2018
Suchstrategie	Lenvatinib OR E7080 OR "E 7080" OR "E-7080" OR "Lenvatinib mesilat" OR Lenvima
Filter	keine
Treffer	81

Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in WHO ICTRP

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	06.07.2018
Suchstrategie	Lenvatinib OR E7080 OR E 7080 OR E-7080 OR lenvatinib mesilat OR Lenvima
Filter	keine
Treffer	116

Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in EU-CTR

Studienregister	EU-CTR
Internetadresse	https://www.clinicaltrialsregister.eu
Datum der Suche	06.07.2018
Suchstrategie	Lenvatinib mesylate OR Lenvatinib OR E7080 OR E-7080 OR "E 7080" OR "lenvatinib mesilat" OR Lenvima
Filter	keine
Treffer	17

Suchstrategie für RCT mit dem zu bewertenden Arzneimittel in PharmNet.Bund

Studienregister	PharmNet.Bund
Internetadresse	https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html
Datum der Suche	06.07.2018
Suchstrategie	Lenvatinib? [Textfelder] OR E7080? [Textfelder] OR E-7080? [Textfelder] OR E 7080? [Textfelder] OR lenvatinib mesilat? [Textfelder] OR Lenvima [Textfelder]
Filter	keine
Treffer	7

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Titel (Autor) der ausgeschlossenen Dokumente	Ausschlussgrund
1	Cheng A. L. et al. Phase 3 trial of lenvatinib (LEN) vs sorafenib (SOR) in first-line treatment of patients (pts) with unresectable hepatocellular carcinoma (uHCC). Oncology Research and Treatment. 2017, 40 (Supplement 3): 211.	Keine Volltextpublikation
2	Desai J. R. et al. Systemic therapy for advanced hepatocellular carcinoma: An update. Journal of Gastrointestinal Oncology. 2017, 8 (2): 243-255.	Kein RCT
3	Finn R. S. et al. A multicenter, open-label, phase 3 trial to compare the efficacy and safety of lenvatinib (E7080) versus sorafenib in first-line treatment of subjects with unresectable hepatocellular carcinoma. Journal of Clinical Oncology. 2014, 32 (15 Supplement 1): No pagination.	Keine Volltextpublikation
4	Finn R. S. et al. Analysis of serum biomarkers (BM) in patients (pts) from a phase 3 study of lenvatinib (LEN) vs sorafenib (SOR) as first-line treatment for unresectable hepatocellular carcinoma (uHCC). Annals of Oncology. 2017, 28 (Supplement 5): v617.	Keine Volltextpublikation
5	Gholam P. The role of Sorafenib in hepatocellular carcinoma. Clinical Advances in Hematology & Oncology. 2015, 13 (4): 232-234.	Keine der empfohlenen Dosierungen von Lenvatinib
6	Gish R. G. und Gholam P. The role of sorafenib in hepatocellular carcinoma. Gastroenterology and Hepatology. 2015, 11 (4): 253-255.	Mehrfachpublikationen ohne relevante Zusatzinformationen
7	Han K. H. et al. Efficacy and safety of lenvatinib for unresectable hepatocellular carcinoma in patients with baseline hepatitis B virus (HBV). Hepatology. 2017, 66 (Supplement 1): 740A-741A.	Keine Volltextpublikation
8	Hudgens S. et al. Time to clinically meaningful worsening in hepatocellular carcinoma patients treated with lenvatinib or sorafenib. Value in Health. 2017, 20 (9): A416.	Keine Volltextpublikation
9	Hudgens S. et al. Detrimental impact of toxicity on quality of life in hepatocellular carcinoma patients treated with lenvatinibor sorafenib. Value in Health. 2017, 20 (9): A411-A412.	Keine Volltextpublikation
10	Tremblay G. et al. Network meta-analysis (NMA) of treatments for unresectable hepatocellular carcinoma (uHCC). Annals of Oncology. 2017, 28 (Supplement 5): v241.	Keine Volltextpublikation

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Studienbezeichnung*	Ausschluss in Teilpopulation A/Ausschlussgrund
1	Clinicaltrials.gov. A Phase I/Ib, Multicenter, Open-Label, Dose Escalation Study of E7080 in Patients With Solid Tumors and in Combination With Temozolomide in Patients With Advanced and/or Metastatic Melanoma. 2015 https://ClinicalTrials.gov/show/NCT00121680	A2
2	Clinicaltrials.gov. An Open Label Phase I Dose Escalation Study Of E7080. 2016 https://ClinicalTrials.gov/show/NCT00121719	A1
3	Clinicaltrials.gov. An Open Label Phase I Dose Escalation Study of E7080 Administered to Patients With Solid Tumors. 2016 https://ClinicalTrials.gov/show/NCT00280397	A2
4	Clinicaltrials.gov. Evaluating the Safety and Efficacy of Oral E7080 in Medullary and Iodine-131 Refractory, Unresectable Differentiated Thyroid Cancers, Stratified by Histology. 2014 https://ClinicalTrials.gov/show/NCT00784303	A2
5	Clinicaltrials.gov. E7080 in Combination With Carboplatin and Paclitaxel in Patients With Non-small Cell Lung Cancer (NSCLC). 2015 https://ClinicalTrials.gov/show/NCT00832819	A1
6	Clinicaltrials.gov. Study of E7080 in Patients With Advanced Hepatocellular Carcinoma (HCC). 2018 https://ClinicalTrials.gov/show/NCT00946153	A5
7	Clinicaltrials.gov. Study of Lenvatinib in Subjects With Advanced Endometrial Cancer and Disease Progression. 2017 https://ClinicalTrials.gov/show/NCT01111461	A1
8	Clinicaltrials.gov. E7080 in Combination With Carboplatin + Gemcitabine Versus Carboplatin + Gemcitabine Alone as Second Line Therapy in Patients With Platinum-Sensitive Recurrent Ovarian Cancer by CA125. 2016 https://ClinicalTrials.gov/show/NCT01133756	A1
9	Clinicaltrials.gov. E7080 (Lenvatinib) in Combination With Dacarbazine Versus Dacarbazine Alone as First Line Therapy in Patients With Stage IV Melanoma. 2016 https://ClinicalTrials.gov/show/NCT01133977	A1

10	Clinicaltrials.gov. A Study of E7080 Alone, and in Combination With Everolimus in Subjects With Unresectable Advanced or Metastatic Renal Cell Carcinoma Following One Prior Vascular Endothelial Growth Factor (VEGF)-Targeted Treatment. 2016 https://ClinicalTrials.gov/show/NCT01136733	A1
11	Clinicaltrials.gov. An Open-Label, 2-Cohort, Multicenter, Study of Lenvatinib in Previously Treated Subjects With Unresectable Stage III or Stage IV Melanoma. 2016 https://ClinicalTrials.gov/show/NCT01136967	A5
12	Clinicaltrials.gov. A Study in Subjects With Recurrent Malignant Glioma. 2016 https://ClinicalTrials.gov/show/NCT01137604	A1
13	Clinicaltrials.gov. E7080 Food Effect Study in Healthy Subjects. 2015 https://ClinicalTrials.gov/show/NCT01240408	A2
14	Clinicaltrials.gov. A Study of E7080 in Subjects With Solid Tumor. 2016 https://ClinicalTrials.gov/show/NCT01268293	A1
15	Clinicaltrials.gov. A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Trial of Lenvatinib (E7080) in 131I-Refractory Differentiated Thyroid Cancer (DTC). 2018 https://ClinicalTrials.gov/show/NCT01321554	A1
16	Clinicaltrials.gov. E7050 in Combination With E7080 in Subjects With Advanced Solid Tumors (Dose Escalation) and in Subjects With Recurrent Glioblastoma or Unresectable Stage III or Stage IV Melanoma After Prior Systemic Therapy (Expansion Cohort and Phase 2). 2016 https://ClinicalTrials.gov/show/NCT01433991	A1
17	Clinicaltrials.gov. A Double-Blind Study in Healthy Volunteers to Assess the Effect of E7080 on the QTc Interval. 2015 https://ClinicalTrials.gov/show/NCT01525394	A2
18	Clinicaltrials.gov. A Study Comparing the Combination of the Best Supportive Care Plus E7080 Versus Best Supportive Care Alone, in Patients With Advanced Lung Cancer or Lung Cancer That Has Spread, Who Have Been Previously Treated, Unsuccessfully, With at Least 2 Different Treatments. 2017 https://ClinicalTrials.gov/show/NCT01529112	A1
19	Clinicaltrials.gov. A Study of E7080 in Subjects With Advanced Thyroid Cancer. 2016 https://ClinicalTrials.gov/show/NCT01728623	A1
20	Clinicaltrials.gov. Study of the Safety and Activity of Lenvatinib (E7080) in Subjects With KIF5B-RET-Positive Adenocarcinoma of the Lung. 2016 https://ClinicalTrials.gov/show/NCT01877083	A1
21	Clinicaltrials.gov. A Pharmacokinetic Study to Assess the Influence of Simultaneous CYP3A4 and P-glycoprotein Inhibition on E7080 Pharmacokinetics Following Single Dose Oral Administration of 5 mg E7080 to Healthy Volunteers. 2015 https://ClinicalTrials.gov/show/NCT02198170	A2
22	Clinicaltrials.gov. A Phase 1, Open-Label, Single-Dose, Pharmacokinetic and Safety Study of E7080 (24 mg) Administered to Subjects With Mild, Moderate, and Severe Renal Impairment and to Healthy Subjects. 2015 https://ClinicalTrials.gov/show/NCT02199379	A5

23	Clinicaltrials.gov. A Pharmacokinetic Study to Assess the Influence of P-glycoprotein Inhibition and Simultaneous CYP3A4 and P-glycoprotein Induction on E7080 Pharmacokinetics Following Single Dose Oral Administration of 24 mg E7080 to Healthy Volunteers. 2015 https://ClinicalTrials.gov/show/NCT02199392	A1
24	Clinicaltrials.gov. An Expanded Access Program With Lenvatinib for the Treatment of Radioiodine-Refractory Differentiated Thyroid Cancer. 2016 https://ClinicalTrials.gov/show/NCT02211222	A1
25	Clinicaltrials.gov. A Pharmacokinetic and Safety Study of E7080 in Subjects With Mild (10 mg), Moderate (10 mg), and Severe Hepatic Impairment (5 mg) and Normal Hepatic Function (10 mg). 2015 https://ClinicalTrials.gov/show/NCT02421042	A2
26	Clinicaltrials.gov. Post-marketing Surveillance of Lenvatinib Mesylate (Lenvima Capsule) in Patients With Unresectable Thyroid Cancer (Study LEN01T). 2017 https://ClinicalTrials.gov/show/NCT02430714	A1
27	Clinicaltrials.gov. Study of Lenvatinib in Children and Adolescents With Refractory or Relapsed Solid Malignancies and Young Adults With Osteosarcoma. 2018 https://ClinicalTrials.gov/show/NCT02432274	A1
28	Clinicaltrials.gov. Study of Lenvatinib in Combination With Everolimus in Patients With Unresectable Advanced or Metastatic Renal Cell Carcinoma (RCC). 2018 https://ClinicalTrials.gov/show/NCT02454478	A1
29	Clinicaltrials.gov. Phase 1b/2 Trial of Lenvatinib (E7080) Plus Pembrolizumab in Subjects With Selected Solid Tumors. 2018 https://ClinicalTrials.gov/show/NCT02501096	A9
30	Clinicaltrials.gov. Neoadjuvant Lenvatinib Combined With Letrozole in Hormone Receptor Positive Breast Cancer. 2016 https://ClinicalTrials.gov/show/NCT02562118	A1
31	Clinicaltrials.gov. A Study to Determine the Metabolism and Elimination of 14C-E7080 in Patients With Advanced Solid Tumors or Lymphomas, Who Are Unsuitable For, or Have Failed, Existing Therapies.. 2017 https://ClinicalTrials.gov/show/NCT02578316	A2
32	Clinicaltrials.gov. Study of Lenvatinib (E7080) in Unresectable Biliary Tract Cancer Who Failed Gemcitabine-based Combination Chemotherapy. 2018 https://ClinicalTrials.gov/show/NCT02579616	A1
33	Clinicaltrials.gov. Tyrosine Kinase Inhibitors (TKI) and Body Composition. 2017 https://ClinicalTrials.gov/show/NCT02592356	A9
34	Clinicaltrials.gov. Eribulin and Lenvatinib in Advanced Solid Tumors. 2018 https://ClinicalTrials.gov/show/NCT02640508	A2
35	Clinicaltrials.gov. A Phase 2 Trial of Lenvatinib for the Treatment of Anaplastic Thyroid Cancer (ATC). 2017 https://ClinicalTrials.gov/show/NCT02657369	A1
36	Clinicaltrials.gov. Lenvatinib Efficacy in Metastatic Neuroendocrine Tumors. 2018 https://ClinicalTrials.gov/show/NCT02678780	A1

37	Clinicaltrials.gov. Phase 1 Study to Determine the Effect of Lenvatinib (E7080) on the Pharmacokinetics of Midazolam in Subjects With Advanced Solid Tumors. 2018 https://ClinicalTrials.gov/show/NCT02686164	A2
38	Clinicaltrials.gov. A Phase 2 Trial of Lenvatinib (E7080) in Subjects With Iodine-131 Refractory Differentiated Thyroid Cancer to Evaluate Whether an Oral Starting Dose of 18 mg Daily Will Provide Comparable Efficacy to a 24 mg Starting Dose, But Have a Better Safety Profile. 2018 https://ClinicalTrials.gov/show/NCT02702388	A1
39	Clinicaltrials.gov. Bioequivalence Study to Evaluate the Impact of Varying Crystalline Polymorph Forms for the Commercial Oral Capsule Formulation of 10-mg Lenvatinib in Healthy Volunteers. 2016 https://ClinicalTrials.gov/show/NCT02723630	A1
40	Clinicaltrials.gov. Phase II Study Assessing the Efficacy and Safety of Lenvatinib for Anaplastic Thyroid Cancer. 2017 https://ClinicalTrials.gov/show/NCT02726503	A1
41	Clinicaltrials.gov. Post-Marketing Surveillance of Lenvima in Korean Patients. 2018 https://ClinicalTrials.gov/show/NCT02764554	A1
42	Clinicaltrials.gov. Testing Lenvatinib in Patients With Adenoid Cystic Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT02780310	A1
43	Clinicaltrials.gov. Lenvatinib and Weekly Paclitaxel for Patients With Recurrent Endometrial or Ovarian Cancer. 2017 https://ClinicalTrials.gov/show/NCT02788708	A1
44	Clinicaltrials.gov. Crossover Study to Evaluate the Relative Bioavailability and Palatability of a Lenvatinib Suspension Compared to the Capsule Formulation in Adult Healthy Volunteers. 2016 https://ClinicalTrials.gov/show/NCT02792829	A2
45	Clinicaltrials.gov. Lenvatinib/Everolimus or Lenvatinib/Pembrolizumab Versus Sunitinib Alone as Treatment of Advanced Renal Cell Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT02811861	A1
46	Clinicaltrials.gov. Study of Lenvatinib in Patients With Advanced Cancer and Aberrations in FGF/FGFR Signaling. 2016 https://ClinicalTrials.gov/show/NCT02846766	A2
47	Clinicaltrials.gov. Lenvatinib in Recurrent and/or Metastatic Adenoid Cystic Carcinomas of the Salivary Glands: ACC-LEN14. 2017 https://ClinicalTrials.gov/show/NCT02860936	A1
48	Clinicaltrials.gov. Lenvatinib and Capecitabine in Patients With Advanced Malignancies. 2017 https://ClinicalTrials.gov/show/NCT02915172	A1
49	Clinicaltrials.gov. A Phase 2 Trial to Evaluate Efficacy and Safety of Lenvatinib in Combination With Everolimus in Subjects With Unresectable Advanced or Metastatic Non Clear Cell Renal Cell Carcinoma (nccRCC) Who Have Not Received Any Chemotherapy for Advanced Disease. 2018 https://ClinicalTrials.gov/show/NCT02915783	A1

50	Clinicaltrials.gov. Capecitabine and Lenvatinib With External Radiation in Rectal Adenocarcinoma. 2018 https://ClinicalTrials.gov/show/NCT02935309	A1
51	Clinicaltrials.gov. Pharmacokinetic Study of E7080/Lenvatinib in Chinese Patients With Unresectable Hepatocellular Carcinoma (HCC). 2018 https://ClinicalTrials.gov/show/NCT02953743	A5
52	Clinicaltrials.gov. A Trial of Lenvatinib (E7080) in Radioiodine (131 I)-Refractory Differentiated Thyroid Cancer in China. 2018 https://ClinicalTrials.gov/show/NCT02966093	A1
53	Clinicaltrials.gov. Lenvatinib and Pembrolizumab in DTC. 2018 https://ClinicalTrials.gov/show/NCT02973997	A1
54	Clinicaltrials.gov. Phase 0 Study in Healthy, Hepatic and Renal Impaired Subjects to Obtain Plasma for Lenvatinib Protein Binding. 2017 https://ClinicalTrials.gov/show/NCT02998775	A9
55	Clinicaltrials.gov. Lenvatinib in Second Line Endometrial Carcinoma. 2016 https://ClinicalTrials.gov/show/NCT03005015	A1
56	Clinicaltrials.gov. Phase 1b Trial of Lenvatinib Plus Pembrolizumab in Subjects With Selected Solid Tumors. 2018 https://ClinicalTrials.gov/show/NCT03006887	A2
57	Clinicaltrials.gov. A Trial of Lenvatinib Plus Pembrolizumab in Subjects With Hepatocellular Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT03006926	A9
58	Clinicaltrials.gov. Lenvatinib in Treating Patients With Metastatic or Advanced Pheochromocytoma or Paraganglioma That Cannot Be Removed by Surgery. 2018 https://ClinicalTrials.gov/show/NCT03008369	A2
59	Clinicaltrials.gov. Pharmacokinetic Study of E7080/Lenvatinib in Chinese Subjects With Solid Tumor. 2018 https://ClinicalTrials.gov/show/NCT03009292	A9
60	Clinicaltrials.gov. Efficacy of Apatinib in Radioactive Iodine-refractory Differentiated Thyroid Cancer. 2017 https://ClinicalTrials.gov/show/NCT03048877	A1
61	Clinicaltrials.gov. UPCC 36315 A Phase II Study Of Everolimus (RAD001) And Lenvatinib (E7080) In Patients With Metastatic Differentiated Thyroid Cancer Who Have Progressed on Lenvatinib Alone. 2017 https://ClinicalTrials.gov/show/NCT03139747	A1
62	Clinicaltrials.gov. Phase II Study of Single Agent Lenvatinib. 2017 https://ClinicalTrials.gov/show/NCT03168074	A1
63	Clinicaltrials.gov. Trial to Assess Safety and Efficacy of Lenvatinib in Combination With Everolimus in Participants With Renal Cell Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT03173560	A1
64	Clinicaltrials.gov. Study of Lenvatinib in Combination With Everolimus in Recurrent and Refractory Pediatric Solid Tumors, Including Central Nervous System Tumors. 2018 https://ClinicalTrials.gov/show/NCT03245151	A5

65	Clinicaltrials.gov. A Study of Lenvatinib, a Multi-targeted Tyrosine Kinase Inhibitor, Combined With Pembrolizumab (PD-1 Inhibitor) for the Treatment of Metastatic Gastroesophageal Cancer Patients Who Have Progressed on First or Subsequent Line Therapies. 2017 https://ClinicalTrials.gov/show/NCT03321630	A2
66	Clinicaltrials.gov. Lenvatinib and Everolimus in Renal Cell Carcinoma (RCC). 2018 https://ClinicalTrials.gov/show/NCT03324373	A1
67	Clinicaltrials.gov. Evaluation of Reporting of Vascular Endothelial Growth Factor and Vascular Endothelial Growth Factor Receptor Inhibitors Associated Cardiovascular Adverse reactioN.. 2018 https://ClinicalTrials.gov/show/NCT03413176	A1
68	Clinicaltrials.gov. Lenvatinib Mesylate and Pembrolizumab in Treating Patients With Metastatic or Recurrent Gastric or Gastroesophageal Cancer. 2018 https://ClinicalTrials.gov/show/NCT03413397	A1
69	Clinicaltrials.gov. A Study of Lenvatinib Plus Nivolumab in Participants With Hepatocellular Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT03418922	A9
70	Clinicaltrials.gov. Non-interventional Study of Cabozantinib in Adults With Advanced Renal Cell Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT03419572	A1
71	Clinicaltrials.gov. Trial to Evaluate Safety and Efficacy of Treatment of Physician Choice (TPC) Following First-Line Treatment of Lenvatinib in Subjects With Unresectable Hepatocellular Carcinoma (uHCC). 2018 https://ClinicalTrials.gov/show/NCT03433703	A9
72	Clinicaltrials.gov. Study to Assess the Long-term Safety of Lenvatinib Monotherapy, a Lenvatinib Combination Regimen, or a Comparator Treatment Arm to Cancer Patients in Eisai Sponsored Lenvatinib Trials. 2018 https://ClinicalTrials.gov/show/NCT03477175	A2
73	Clinicaltrials.gov. Phase 1b of ASLAN001 (Varlitinib) in Patients With Advanced/ Metastatic Hepatocellular Carcinoma (HCC). 2018 https://ClinicalTrials.gov/show/NCT03499626	A9
74	Clinicaltrials.gov. Lenvatinib and Iodine Therapy in Treating Patients With Radioactive Iodine-Sensitive Differentiated Thyroid Cancer. 2018 https://ClinicalTrials.gov/show/NCT03506048	A1
75	Clinicaltrials.gov. A Study of Biomarker-Directed, Pembrolizumab (MK-3475) Based Combination Therapy for Advanced Non-Small Cell Lung Cancer (MK-3475-495). 2018 https://ClinicalTrials.gov/show/NCT03516981	A1
76	Clinicaltrials.gov. Lenvatinib in Combination With Pembrolizumab Versus Treatment of Physician's Choice in Participants With Advanced Endometrial Cancer (MK-3475-775/E7080-G000-309 Per Merck Standard Convention [KEYNOTE-775]). 2018 https://ClinicalTrials.gov/show/NCT03517449	A2
77	Clinicaltrials.gov. Testing Lenvatinib and Cetuximab in Patients With Advanced Head and Neck Squamous Cell Carcinoma and Cutaneous Squamous Cell Carcinoma. 2018 https://ClinicalTrials.gov/show/NCT03524326	A1

78	Clinicaltrials.gov. Lenvatinib and Eribulin in Advanced Soft Tissue Sarcoma. 2018 https://ClinicalTrials.gov/show/NCT03526679	A2
79	Clinicaltrials.gov. Expanded Access Program With Lenvatinib for the Treatment of Differentiated Thyroid Cancer in Brazil. 2018 https://ClinicalTrials.gov/show/NCT03533361	A1
80	Clinicaltrials.gov. A Study to Evaluate the Safety and Efficacy of Lenvatinib in Participants With Refractory Differentiated Thyroid Cancer. 2018 https://ClinicalTrials.gov/show/NCT03573960	A1
81	EU CTR. Phase II, Multicenter, Open-label, Single Arm Trial to Evaluate the Safety and Efficacy of Oral E7080 in Medullary and Iodine-131 Refractory, Unresectable Differentiated Thyroid Cancers, Stratified.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2007-005933-12	A9
82	EU CTR. An Open-Label, Single-Arm, Multicenter Phase 2 Study of E7080 in Subjects with Advanced Endometrial Cancer and Disease Progression Following First-Line Chemotherapy. 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-016858-41	A1
83	EU CTR. An open-label, multicenter Phase Ib/2 study of E7080 alone, and in combination with everolimus in subjects with unresectable advanced or metastatic renal cell carcinoma following one prior VEGF-tar.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-019484-10	A1
84	EU CTR. A Multicenter, Randomized, Double- Blind, Placebo-Controlled, Phase 3 Trial of Lenvatinib (E7080) in 131I- Refractory Differentiated Thyroid Cancer. 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-023783-41	A1
85	EU CTR. A Phase II, Randomized, Double-Blind, Placebo-Controlled Study of Oral E7080 in Addition to Best Supportive Care (BSC) versus BSC Alone in Patients with Locally Advanced or Metastatic Non-Squamous .... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-002347-10	A1
86	EU CTR. Phase 1/2 Study of Lenvatinib in Children and Adolescents With Refractory or Relapsed Solid Malignancies and Young Adults with Osteosarcoma . 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-005534-38	A1
87	EU CTR. A Multicenter, Randomized, Double-blind Phase 2 Trial of Lenvatinib (E7080) in Subjects with 131I Refractory Differentiated Thyroid Cancer to Evaluate Whether an Oral Starting Dose of 20 mg or 14 m.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-005199-27	A9

88	EU CTR. This is an Open-Label, Single-Arm, Multicenter, Phase 2 Trial of Lenvatinib for the Treatment of Anaplastic Thyroid Cancer (ATC).. 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-001929-17	A1
89	EU CTR. The Drug Rediscovery Protocol (DRUP trial) A Dutch National Study on behalf of the Center for Personalized Cancer Treatment (CPCT) to Facilitate Patient Access to Commercially Available, Targeted A.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-004398-33	A9
90	EU CTR. An Open-Label, 2-Cohort, Multicenter, Phase 2 Study of E7080 in Previously Treated Subjects With Unresectable Stage III or Stage IV Melanoma. 0000 https://portal.dimdi.de/data/ctr/O-2712050-1-0-634F93-20170627145312.pdf	A1
91	EU CTR. A Multicenter, Open-label, Randomized, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib in Combination with Everolimus or Pembrolizumab Versus Sunitinib Alone in First-Line Treatment .... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-000916-14	A1
92	EU CTR. A Randomized, Double-blind, Phase 2 Trial to Assess Safety and Efficacy of Lenvatinib at Two Different Starting Doses (18 mg vs. 14 mg QD) in Combination with Everolimus (5 mg QD) in Renal Cell Car.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-002778-11	A1
93	EU CTR. A Multicenter, Open-Label Phase 1b/2 Trial of Lenvatinib (E7080) Plus Pembrolizumab in Subjects With Selected Solid Tumors. 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000300-26	A9
94	EU CTR. An open-label, multi-center, roll-over study to assess long term safety of lenvatinib monotherapy or lenvatinib combination regimen or comparator treatment arm to cancer patients in Eisai sponsored.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003668-11	A9
95	EU CTR. A Multicenter, Open-label, Randomized, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib in Combination with Pembrolizumab Versus Treatment of Physician's Choice in Participants with A.... 0000 https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-004387-35	A1
96	PharmNet.Bund. An Open-Label, 2-Cohort, Multicenter, Phase 2 Study of E7080 in Previously Treated Subjects With Unresectable Stage III or Stage IV Melanoma. 0000 WWW.PHARMNET-BUND.DE	A1
97	PharmNet.Bund. Enhancing Radioiodine Incorporation into Radio Iodine Refractory Thyroid Cancers with MAPK Inhibition: A single center pilot study. 0000 WWW.PHARMNET-BUND.DE	A1

98	WHO ICTRP. Sorafenib in Treating Patients With Ovarian Epithelial Cancer, Fallopian Tube Cancer, or Peritoneal Cancer in at Least the Second Remission. 2017 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00522301	A1
99	WHO ICTRP. O6-Benzylguanine and Topical Carmustine in Treating Patients With Early-Stage IA-IIA Cutaneous T-Cell Lymphoma. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00961220	A2
100	WHO ICTRP. The Drug Rediscovery Protocol (DRUP trial). 2017 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-004398-33-NL	A1
101	WHO ICTRP. An Open-Label, Multicenter, Randomized, Phase Ib/II Study of E7080 in Combination with Dacarbazine versus Dacarbazine Alone as First Line Therapy in Patients with Stage IV Melanoma.. 2012 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-016049-24-GB	A1
102	WHO ICTRP. An Open-Label, Multicenter, Randomized, Phase Ib/II Study of E7080 in Combination with Carboplatin + Gemcitabine versus Carboplatin + Gemcitabine Alone as Second Line Therapy in Patients with Platinum-Sensitive Recurrent Ovarian Cancer by CA125.. 2012 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-016050-41-PL	A1
103	WHO ICTRP. TRIAL TO ASSESS THE EFFICACY OF LENVATINIB IN METASTATIC NEUROENDOCRINE TUMORS (TALENT STUDY).?. 2015 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-001467-39-ES	A1
104	WHO ICTRP. Analysis of therapeutic response onset time in thyroid cancer patients treated by lenvatinib. 2016 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000019484	A1
105	WHO ICTRP. Phase II study assessing the efficacy and safety of lenvatinib for anaplastic thyroid cancer (HOPE). 2016 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000020773	A1
106	WHO ICTRP. Evaluation of early prediction using 18F-FDG PET/CT for patients with Iodine-131 refractory differentiated thyroid carcinoma treated with lenvatinib. 2016 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000022592	A1
107	WHO ICTRP. A phase I study of lenvatinib for unresectable adenoid cystic carcinoma of the salivary gland. 2017 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000021080	A1
108	WHO ICTRP. Phase II study assessing the efficacy and safety of lenvatinib for anaplastic thyroid cancer. 2016 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000021512	A9
109	WHO ICTRP. A prospective study of efficacy and safety of lenvatinib in patients with RAI-refractory differentiated thyroid cancer patients. 2016 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000021757	A1

110	WHO ICTRP. Phase II trial of Lenvatinib in patients with metastatic or . 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000026777	A1
111	WHO ICTRP. An Open-Label, Expanded Access Program With Lenvatinib for the First Line Treatment of Unresectable Hepatocellular Carcinoma. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-JapicCTI-173761	A9
112	WHO ICTRP. Phase II study exploring reduction method of Lenvatinib for the patients with advanced hepatocellular carcinoma. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000032400	A9
113	WHO ICTRP. A phase II trial of lenvatinib in patients with metastatic colorectal cancer after standard chemotherapy. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000023446	A1
114	WHO ICTRP. A phase II trial of lenvatinib in patients with metastatic colorectal cancer after standard chemotherapy. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-JMA-IIA00261	A1
115	WHO ICTRP. Combination of molecularly-targeted therapy with stereotactic body radiotherapy in hepatocellular carcinoma with major vascular invasion. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000030718	A5
116	WHO ICTRP. Cohort Study to Explore the Therapeutic Effect of Lenvatinib on Differentiated Thyroid Cancer. 2017 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000022243	A1
117	WHO ICTRP. A clinical study in management of Cerebral Palsy with Shashtika Shali Pinda Sweda and Yoga Basti. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2016/07/007080	A1
118	WHO ICTRP. A phase 2 study of E7080 in subjects with advanced thyroid cancer. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-JapicCTI-122009	A1
119	WHO ICTRP. Retrospective study for Lenvatinib for unresectable thyroid cancers. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-UMIN000030969	A5
120	WHO ICTRP. Post Marketing Phase IV study to assess the Safety and Efficacy of Lenvatinib in Subjects with Thyroid Cancer.. 2018 http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2018/04/012994	A1
* Studien, die in mehr als einem Register identifiziert wurden, werden einmalig aufgeführt.		

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-84 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.3.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-84 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-84 (Anhang): Studiendesign und -methodik für Studie 304

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2b	Genaue Ziele, Fragestellung und Hypothesen	**Ziele:** Die Phase-III-Studie wurde durchgeführt zur Evaluierung der Wirksamkeit und Sicherheit von Lenvatinib im Vergleich zu Sorafenib als systemische Erstlinientherapie bei Patienten mit einem inoperablem Leberzellkarzinom (Hepatocellular carcinoma (HCC)).
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei der vorliegenden Studie handelt es sich um eine multizentrische, randomisierte, kontrollierte Open-Label Phase-III-Studie.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Drei Änderungen des Studienprotokolls wurden bis zum Datenschnitt für die primäre Analyse durchgeführt. **Änderung 1 (06. Januar 2014):** • Revision des Einschlusskriteriums #7, um die Definition einer adäquaten Nierenfunktion anhand des Kreatinin-Clearance-Wert von > 30 auf > 40 mL/min zu ändern. Dabei wird der Wert mit Hilfe der Cockcroft-Gault Formel berechnet. • Revision des Einschlusskriteriums #9, um die erlaubte Anzahl an Antihypertonika, die für eine adäquate Kontrolle des Blutdrucks eines Patienten benötigt wird, von 0 auf 1 und dann auf bis zu 3 zu erhöhen. • Klarstellung des Ausschlusskriteriums #2, um zu spezifizieren, dass Patienten, die eine Anti-VEGF Therapie, nicht nur Sorafenib, erhalten haben, aus der Studie ausgeschlossen wurden. • Klarstellung des Ausschlusskriteriums #8, um zu spezifizieren, dass Patienten, die Antikoagulantien einnahmen, die ein International Normalized Ratio (INR) Monitoring benötigen, aus der Studie ausgeschlossen wurden.

Item[a]	Charakteristikum	Studieninformation
		• Klarstellung des Ausschlusskriteriums #10, dass Patienten, die eine Behandlung einer aktiven Magenblutung oder von Krampfadern benötigten oder bei denen ein chirurgischer Eingriff notwendig war, um der Blutung vorzubeugen, aus der Studie ausgeschlossen wurden. • Aufnahme des Ausschlusskriteriums #12, um zu spezifizieren, dass Patienten, deren einzige Zielläsion im Knochen war, aus der Studie ausgeschlossen wurden. • Revision des Ausschlusskriteriums #20, um zu spezifizieren, dass Patienten mit einer Drogen- oder Alkoholabhängigkeit oder einem -missbrauch in den sechs Monaten vor Studienbeginn, aus der Studie ausgeschlossen wurden. • Wie in der EU Voluntary Harmonisation Procedure (VHP) angefordert, wurden die folgenden Maßnahmen zur Überwachung der Sicherheit in das Protokoll aufgenommen: o regelmäßige Kontrolle der Leberfunktionstests und Guidelines für das Umgehen mit einer verminderten Leberfunktion o Kontrolle der Symptome venöser Thromboembolien und Guidelines für den Umgang mit diesen Ereignissen • Revision der sekundären Zielgrößen, um die Ergebnisse des EORTC QLQ-C30 Fragebogens zur Feststellung des Einflusses der Behandlung auf die gesundheitsbezogene Lebensqualität (engl. health-related quality of life, HRQoL)der Patienten, die mit Lenvatinib versus Sorafenib behandelt wurden, zu erfassen; das Ausfüllen der HRQoL Fragebögen wurde in den Auswertungsplan aufgenommen und die statistischen Methoden wurden angepasst. • Löschung der Beurteilung der Health State Utilities, die mit der objektiven Response, gemessen mit Hilfe des EQ-5D Fragebogens, in Verbindung stehen. • Aufnahme einer explorativen Zielgröße und eines explorativen Endpunkts, um die Blut und Tumor Biomarker, die mit dem klinischen Outcome korreliert sein könnten, zu beurteilen; die Entnahme von festem Tumorgewebe und Biomarker Blutproben wurden in den Auswertungsplan aufgenommen und die statistischen Methoden wurden angepasst. • Klarstellung im Protokoll, dass wenn ein Patient die Studie vor der ersten Tumorbewertung, die für Woche 8 geplant war, abgebrochen hatte, diese Bewertung so nah wie möglich an Woche 8 aber vor dem Beginn einer weiteren Anti-Krebs-Therapie erfolgen sollte. • Klarstellung im Protokoll, dass eine Toxizität von Grad 2 vom Patient und vom Prüfarzt als tolerierbar

Item[a]	Charakteristikum	Studieninformation
		oder nicht-tolerierbar festgelegt wird; aufgrund tolerierbarer Toxizität vom Grad 2 wurde die Studienmedikation mit oder ohne Unterbrechung reduziert. Die Unterbrechung der Studienmedikation war für Toxizität vom Grad 3 vorgeschrieben. • Klarstellung im Protokoll, dass adipöse Patienten mit Gewichtsverlust nicht zum Baseline oder Gewichtsverlust vom Grad 1 zurückkehren mussten, um die Behandlung mit Lenvatinib wieder aufzunehmen; jedoch durfte der Gewichtsverlust nicht länger als eine Woche anhalten, Lenvatinib sollte mit einer niedrigeren Dosis wieder aufgenommen werden und ein normaler Body Mass Index sollte für weitere Dosisreduktionen herangezogen werden. • Aufnahme einer Dosierungsanleitung für Lenvatinib und Guidelines zum Bluthochdruck Management für Patienten mit dem Risiko eine hypertensive Krise zu entwickeln oder mit signifikanten Risikofaktoren für schwere Komplikationen bei unkontrolliertem Bluthochdruck. • Klarstellung, dass Patienten, die mit Sorafenib behandelt werden, zweimal täglich zwei 200 mg Tabletten (400 mg) einnehmen sollten und dass bei Patienten mit einer Toxizität, die auf Sorafenib zurückzuführen ist, gemäß den Sorafenib Kennzeichnungsrichtlinien vorgegangen werden soll. • Klarstellung, dass die folgenden Begleittherapien erlaubt sind o palliative Strahlentherapie von bis zu zwei schmerzhaften, schon bestehenden,Nichtziel-Knochenmetastasen o Antivirale Therapie für Patienten mit Hepatitis B, die diese Therapie bereits erhalten o Behandlung mit Antikoagulantien, die kein INR Monitoring benötigen • Klarstellung, dass eine gesamte Blutprobe für pharmakogenomische Analysen nur von Patienten, die mit Lenvatinib behandelt werden, entnommen wird. • Klarstellung, dass Patienten mindestens eine Dosis der Studienmedikation erhalten haben mussten, um in das Safety Analysis Set eingeschlossen zu werden; die Voraussetzung, dass die Patienten mindestens eine post-Baseline Bewertung haben mussten, wurde entfernt. • Klarstellung, dass die Studieneinschreibung in China auf ungefähr 200 Patienten beschränkt ist. • Aufnahme der Voraussetzung, dass Prüfärzte Filme der Echokardiogramme und der multiple gated

Item[a]	Charakteristikum	Studieninformation
		acquisition scan (MUGA) Scans im zentralen Datenspeicher ablegen müssen. • Klarstellung, dass alle SUE, unabhängig davon ob sie in Zusammenhang mit der Studie stehen, auf einem vollständigen SUE Formblatt per E-Mail oder Fax so bald wie möglich, aber nicht später als 24 Stunden nach dem Zeitpunkt an dem der Prüfarzt das Ereignis festgestellt hat, berichtet werden müssen. • Änderung des Intervalls für die Nachbeobachtung der Überlebensdaten von alle zwei Monate auf alle 12 Wochen. **Änderung 2 (09. April 2014):** • Für den Fall, dass eine Progression des HCC zu einem unerwünschten Ereignis führt, sollte dies als unerwünschtes Ereignis, das während der Behandlung auftrat (Treatment emergent adverse event, TEAE) gemeldet werden. Die Progression des HCC alleine wird nicht als TEAE gemeldet. • Revision des Ausschlusskriteriums Nr. 10, um zu verdeutlichen, dass Patienten, die eine aktive interventionelle Behandlung von Magen- oder Ösophagusvarizen innerhalb von 28 Tagen vor Randomisierung erhalten haben, ausgeschlossen werden. • Klärung des Toxizitätsmanagements für mit Sorafenib behandelte Patienten mit Verlängerung des frequenzkorrigierten QT-Zeit-Intervalls (Q-T Corrected, QTc). (≥ 501 ms) und Raten zur Vorsicht bei zeitgleicher Verwendung von Sorafenib und Arzneimitteln, die bekannt dafür sind, das QTc-Intervall zu verlängern. **Änderung 3 (16. Januar 2015):** • Informationen über das Auftreten, die Diagnose und das Management des Posterioren Reversiblen Enzephalopathiesyndroms (PRES) bei mit Lenvatinib behandelten Patienten wurden aufgenommen. • Da das Auftreten von PRES in die Lenvatinib Broschüre des Prüfarztes als „erwartete“ Nebenwirkung aufgenommen wurde, waren weitere Berichte an die Regulierungsbehörden über PRES nicht mehr notwendig.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** 1. Patienten mit einer bestätigten Diagnose eines inoperablen HCC: o histologisch oder zytologisch bestätigte Diagnose eines HCC o klinisch bestätigte Diagnose eines HCC gemäß den Kriterien der American Association for the Study of Liver Diseases (AASLD), einschließlich Zirrhosen jeder

Item[a]	Charakteristikum	Studieninformation
		Etiologie, oder mit chronischer Hepatitis B (HBV) oder C (HCV)-Infektion 2. Patienten mit mindestens einer messbaren Zielläsion gemäß der modified Response Evaluation Criteria in Solid Tumors (mRECIST): o hepatische Läsion: I. die in mindestens einer Dimension ≥ 1 cm misst (lebensfähiger Tumor für typische und längster Durchmesser für atypische) II. die für wiederholte Messungen geeignet ist o nicht-hepatische Läsion: I. Lymphknoten Läsion, die in mindestens einer Dimension ≥ 1,5 cm in der kurzen Achse misst, mit Ausnahme der Porta Hepatis Lymphknoten Läsion, die ≥ 2,0 cm in der kurzen Achse misst II. Nicht-nodale Läsion, dessen längster Durchmesser ≥ 1,0 cm beträgt 3. Patienten, die in Stadium B (nicht geeignet für TACE) oder Stadium C gemäß dem Barcelona Clinic Liver Cancer (BCLC) Staging-System kategorisiert wurden. 4. Patienten mit adäquater Knochenmarkfunktion, definiert durch die Werte o Absolute Neutrophilenzahl (ANC) $\geq 1{,}5 \times 10^9$/L o Hämoglobin (Hb) ≥ 8,5 g/dL o Thrombozytenzahl $\geq 75 \times 10^9$/L 5. Patienten mit adäquater Leberfunktion, definiert durch die Werte o Albumin ≥ 2,8 g/dL o Bilirubin ≤ 3,0 mg/dL o Aspartat-Aminotransferase (AST), alkalische Phosphatase (ALP) und Alanin-Aminotransferase (ALT) ≤ 5 × Obergrenze für Normalwerte (engl. upper linit of normal, ULN) 6. Patienten mit adäquater Blutgerinnungsfunktion, definiert durch den Blutgerinnungswert INR (International Normalized Ratio) ≤ 2,3. 7. Patienten mit adäquater Nierenfunktion definiert durch den Kreatinin-Clearance-Wert > 40 mL/min, berechnet mit Hilfe der Cockcroft-Gault Formel. 8. Patienten mit adäquater Bauchspeicheldrüsenfunktion, definiert durch den Amylase- und Lipase-Wert ≤ 1,5 × ULN.

Item[a]	Charakteristikum	Studieninformation
		9. Patienten mit einem adäquat mit bis zu drei Antihypertonika kontrollierten Blutdruck (engl. blood pressure, BP), definiert durch den BP-Wert ≤ 150/90 mm Hg bei Screening und keiner Änderung der Antihypertonika in der ersten Woche vor Zyklus 1/Tag 1. 10. Patienten mit Child-Pugh A. 11. Patienten mit einem Eastern Cooperative Oncology Group Performance Status (ECOG PS) von 0 oder 1. 12. Patienten mit einer Lebenserwartung von ≥ 12 Wochen nach Beginn der Einnahme der Studienmedikation. 13. Männliche und weibliche Patienten im Alter von ≥ 18 Jahren (oder erwachsene Patienten jeglichen Alters ≥ 18 Jahren, gemäß der Festlegung in der länderspezifischen Gesetzgebung). 14. Weibliche Patientinnen im gebärfähigen Alter durften zu Baseline nicht stillen oder schwanger sein (d.h. negativer beta-humaner Choriongonadotropin [ß-hCG] -Test mit einer minimalen Empfindlichkeit von 25 IU/L oder äquivalenter Einheit zu ß-hCG). Eine separate Bewertung zu Baseline war erforderlich, falls der Schwangerschaftstest mit negativem Ergebnis mehr als 72 Stunden vor der Verabreichung der ersten Dosis der Studienmedikation durchgeführt wurde. 15. Weibliche Patientinnen wurden als gebärfähig erachtet, es sei denn sie waren postmenopausal (Amenorrhoe für mindestens 12 aufeinanderfolgende Monate in der entsprechenden Altersgruppe und ohne andere bekannte oder vermutete Ursachen) oder sie wurden operativ sterilisiert (d.h. bilaterale Tubenligatur, totale Hysterektomie oder bilaterale Oophorektomie mindestens ein Monat vor der Dosierung). 16. Weibliche Patientinnen im gebärfähigen Alter durften keinen ungeschützten Geschlechtsverkehr innerhalb der 30 Tage vor Studieneintritt haben oder musste zustimmen über die gesamte Studiendauer und 30 Tage nach Einstellung der Studienmedikation eine hoch effektive Verhütungsmethode zu verwenden (z.B. totale Abstinenz, ein Intrauterinsystem, eine Doppelbarrierenmethode [wie Kondom plus Diaphragma mit Spermizid], ein Verhütungsimplantat, ein orales Verhütungsmittel oder ein vasektomierter Partner mit bestätigter Azoospermie). Wenn die Patientin aktuell abstinent war, musste sie zustimmen, eine Doppelbarrierenmethode, wie oben beschrieben, zu verwenden, falls sie während der Studie oder 30 Tage nach Einstellung der Studienmedikation wieder sexuell aktiv wird. Patientinnen, die ein hormonelles Verhütungsmittel verwendeten, mussten für mindestens vier Wochen vor Dosierung eine stabile

Item[a]	Charakteristikum	Studieninformation
		Dosis des gleichen hormonellen Verhütungsmittel einnehmen und mussten das gleiche hormonelle Verhütungsmittel während der Studie und 30 Tage nach Einstellung der Studienmedikation verwenden. 17. Männliche Patienten mussten eine erfolgreiche Vasektomie gehabt haben (bestätigte Azoospermie) oder ihre weibliche Partnerin musste ein oben erwähntes Kriterium erfüllen (d.h. nicht gebärfähig sein oder eine hoch effektive Verhütungsmethoden während der Studiendauer und 30 Tage nach Einstellung der Studienmedikation verwenden). 18. Patienten mussten eine schriftliche Einverständniserklärung abgeben. 19. Patienten mussten bereit und in der Lage sein alle Aspekte des Protokolls zu erfüllen. **Ausschlusskriterien:** 1. Patienten mit einem bildgebenden Befund für HCC, der einem der folgenden Punkte entspricht: o HCC mit ≥ 50% Leberbesetzung o klare Invasion in den Gallengang o Invasion der Pfortader am Hauptast 2. Patienten, die eine systemische Chemotherapie, einschließlich der Anti-VEGF Therapie, oder eine systemische Anti-Krebs Medikation, einschließlich Lenvatinib bei fortgeschrittenem / inoperablem HCC, erhalten haben. 3. Patienten, die eine Anti-Krebs-Therapie (einschließlich Operation, perkutane Ethanol-Injektion, Radiofrequenzablation, Transarterielle (Chemo-) Embolisation, hepatische intraarterielle Chemotherapie, biologisch, Immuntherapie, hormonell oder Radiotherapie) oder eine blutqualitätfördernde Medikation (einschließlich Bluttransfusion, Blutprodukte oder Mittel, die die Blutzellenproduktion stimuliert, z.B. G-CSF [granulocyte colony-stimulating factor]) innerhalb 28 Tage nach Randomisierung erhalten haben. 4. Patienten, die sich nicht von der Toxizität einer vorangegangenen Anti-Krebs-Therapie erholt haben, außer Alopezie und Unfruchtbarkeit; dabei galt ein Patient als von der Toxizität erholt, wenn ein Schweregrad < 2 gemäß den Common Terminology Criteria for Adverse Events (CTCAE) v4.0 vorlag. 5. Patienten mit signifikanter Herz-Kreislauf-Erkrankung: o Anamnese einer kongestiven Herzinsuffizienz mit einem von der New York Heart Association (NYHA) definierten Schweregrad größer als Stadium II o instabile Angina Pectoris

Item[a]	Charakteristikum	Studieninformation
		○ Myokardinfarkt oder Schlaganfall innerhalb von sechs Monaten nach dem Erhalt der ersten Dosis der Studienmedikation ○ Herzrhythmusstörungen, die beim Screening eine medizinische Behandlung erforderten 6. Patienten mit einer QTc-Intervall-Verlängerung auf > 480ms. 7. Patienten mit einer gastrointestinalen Malabsorption oder einer anderen Erkrankung, die nach Ansicht des Prüfarztes die Absorption von Lenvatinib beeinflussen könnte. 8. Patienten mit Blutungen oder thrombotischen Störungen oder die Antikoagulantien einnahmen, die ein (INR Monitoring benötigen, z.B. Warfarin oder ähnliche Medikationen. Die Behandlung mit niedermolekularen Heparin- oder Faktor X-Inhibitoren, die kein INR Monitoring erfordern, war zulässig. Antithrombozytenagenten waren während der gesamten Studie verboten. 9. Patienten mit einer gastrointestinalen Blutung oder einer aktiven Hämoptyse (mindestens 0,5 Teelöffel hellrotes Blut) innerhalb von 28 Tagen vor Randomisierung. 10. Patienten mit Magen- oder Ösophagusvarizen, die eine aktive interventionelle Behandlung innerhalb von 28 Tagen vor Randomisierung erforderten. Die Prophylaxe mit pharmakologischer Therapie (z.B. nichtselektive Beta-Blocker) war erlaubt. 11. Patienten mit aktiven malignen Tumoren (mit Ausnahme von HCC oder endbehandelten Melanomen in-situ, basal oder Plattenepithelkarzinomen der Haut, oder Karzinomen in-situ des Gebärmutterhalses) in den vergangenen 36 Monaten. 12. Patienten, deren einzige Zielläsion im Knochen war. 13. Patienten mit Meningealer Karzinomatose. 14. Patienten mit der Anamnese von oder mit bestehenden Hirnmetastasen oder mit subduralen Metastasen. 15. Patienten mit einem Ergebnis von > 1 + beim Test auf Proteinurie, deren Urin für eine quantitative Analyse der Urinproteine für 24 Stunden gesammelt und dann einen Urinproteinwert von ≥ 1 g / 24-Stunden aufwiesen. 16. Patienten mit einem chirurgischen arterio-portalen oder arterio-venösen Shunt. 17. Patienten mit jeglichen medizinischen oder anderen Zuständen, die nach Auffassung des Prüfarztes die Teilnahme des Patienten an einer klinischen Studie ausschließt.

Item[a]	Charakteristikum	Studieninformation
		18. Patienten mit bekannter Intoleranz gegenüber Lenvatinib oder Sorafenib (oder gegenüber einer der Hilfsstoffe). 19. Patienten, die positiv auf das Humane Immundefizienz-Virus (HIV) getestet wurden oder Patienten mit einer aktiven Infektion, die eine Behandlung benötigte (außer Hepatitis-Viren). 20. Patienten mit einer Drogen- oder Alkoholabhängigkeit oder einem -missbrauch in den sechs Monaten vor Studienbeginn. 21. Patienten, die nicht mit einer dreiphasigen CT der Leber oder einer dreiphasigen MRI der Leber aufgrund einer Allergie oder einer anderen Kontraindikation für CT und MRI Kontrastmittel untersucht werden können. 22. Patienten mit einer großen Operation, die innerhalb von drei Wochen vor Randomisierung oder während der Studie stattfand. 23. Patienten mit einem Lebertransplantat.
4b	Studienorganisation und Ort der Studiendurchführung	Die multizentrische Studie wurde in 183 Zentren in Asien, Nordamerika, der Europäischen Union, Russland und Israel durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	**Interventionsgruppe** Lenvatinib wurde einmal täglich in 28-Tage Zyklen in folgenden Dosierungen oral verabreicht (immer zur gleichen Tageszeit konsistent mit oder ohne Essen): • Patienten mit einem Körpergewicht ≥ 60 kg zu Baseline: 12 mg Lenvatinib (drei 4 mg-Kapseln) • Patienten mit einem Körpergewicht < 60 kg zu Baseline: 8 mg Lenvatinib (zwei 4 mg-Kapseln) **Kontrollgruppe** 400 mg Sorafenib (NEXAVAR®) (zwei 200 mg-Tabletten) wurde zweimal täglich oral verabreicht (immer zur gleichen Tageszeit konsistent mit oder ohne Essen).
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	**Primäre Zielgröße:** • Gesamtüberleben (engl. Overall Survival, OS), definiert als die Zeit vom Zeitpunkt der Randomisierung bis zum Eintreten des Todes aufgrund jeglicher Ursache. Lost-to-follow-up-Patienten wurden zum letzten Zeitpunkt, an dem sie bekanntermaßen am Leben waren, zensiert. Überlebende Patienten wurden zum Zeitpunkt des Datenschnitts zensiert. **Sekundäre Zielgrößen:** • Progressionsfreies Überleben (engl. Progression Free Survival, PFS), definiert als Zeit vom Zeitpunkt der Randomisierung bis zum Zeitpunkt der ersten Dokumentation einer Krankheitsprogression oder bis zum Zeitpunkt des Versterbens, bestimmt gemäß

		mRECIST. Die Regeln für die Zensierung wurden im Statistischen Analyseplan (engl. statistical analysis plan, SAP) definiert und folgen den Richtlinien der FDA. Patienten, die die Therapie aufgrund einer Krankheitsprogression beendeten oder bei denen keine radiologische Bestätigung für die Krankheitsprogression vorlag, wurden zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert. Patienten, die eine neue Therapie anfingen, verstarben oder bei denen eine Progression nach mehr als einer verpassten Visite oder nach 28 Tagen der letzten Dosis der Studienmedikation auftrat, wurden ebenfalls zum letzten Zeitpunkt, zu dem eine radiologische Untersuchung stattfand, zensiert. Die Untersuchungen wurden vom entsprechend geschulten, qualifizierten Personal durchgeführt und die Ergebnisse des Prüfarztes in der Clinical Database (CRF) dokumentiert und an das Imaging Core Laboratory (ICL) Parexel International übermittelt. • Zeit zur Progression (engl. Time to Progression, TTP), definiert als Zeit vom Zeitpunkt der Randomisierung bis zum Zeitpunkt der ersten Dokumentation einer Krankheitsprogression, bestimmt gemäß mRECIST. Die Regeln für die Zensierung entsprechen denen, die unter dem Endpunkt PFS aufgeführt wurden mit der Ausnahme, dass Tod zensiert wurde. Die Untersuchungen wurden von entsprechend qualifiziertem Personal durchgeführt und die Ergebnisse des Prüfarztes in der Clinical Database (CRF) dokumentiert und an das ICL Parexel International übermittelt. • Objektive Ansprechrate (engl. Objective Response Rate, ORR), definiert durch den Anteil der Patienten mit vollständigem Tumoransprechen (engl. Complete Response, CR) oder teilweisem Tumoransprechen (engl. Partial Response, PR) als insgesamt bestes Tumoransprechen, bestimmt gemäß mRECIST. Die Regeln für die Zensierung entsprechen denen, die unter dem Endpunkt PFS aufgeführt wurden. Die Untersuchungen wurden von entsprechend qualifiziertem Personal durchgeführt und die Ergebnisse des Prüfarztes in der Clinical Database (CRF) dokumentiert und an das ICL Parexel International übermittelt. • HRQoL, gemessen mithilfe des EORTC QLQ-C30 (European Organization for Research and Treatment of Cancer - Quality of Life Questionnaire), dem EORTC HCC-spezifischen Fragebogen (EORTC Supplement for Patients with Hepatocellular Carcinoma, EORTC QLQ-HCC18) und dem generischen Instrument EQ-5D (EuroQoL Five Dimension Scale). • Pharmakokinetische Expositionsparameter von Lenvatinib (Plasma). • Sicherheit und Verträglichkeit von Lenvatinib

Item[a]	Charakteristikum	Studieninformation
		Explorative Zielgrößen: • Disease Control Rate (DCR), definiert durch den Anteil der Patienten mit CR, PR oder stabilem Krankheitsverlauf (Stable Disease, SD; SD ≥ 7 Wochen nach Randomisierung) als insgesamt bestes Tumoransprechen. • Clinical Benefit Rate (CBR), definiert durch den Anteil der Patienten mit CR, PR oder mit langfristig stabilem Krankheitsverlauf (Stable Disease; SD ≥ 23 Wochen nach Randomisierung) als insgesamt bestes Tumoransprechen. • Exploratorische Biomarker Analysen, und ihre Korrelation mit den Wirksamkeitsendpunkten dieser Studie.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlplanung beruhte primär auf der erforderlichen Anzahl von Ereignissen für die Aufdeckung der Nicht-Unterlegenheit (Schritt 1) und Überlegenheit (Schritt 2) bezüglich OS von Lenvatinib gegenüber Sorafenib. Die Schätzung der benötigten Anzahl von Ereignissen beruhte auf folgenden Annahmen: • Für OS wurde eine exponentielle Verteilung angenommen. Das geschätzte mediane OS unter Sorafenib betrug etwa 10 Monate. Eine Verbesserung um 2,5 Monate wurde aus dem angestrebten Hazard Ratio (HR) von 0,8 abgeleitet, was einen deutlichen klinischen Nutzen darstellen würde. • Unter Verwendung eines Nicht-Unterlegenheits-Tests des logarithmierten HR von OS nach der 95%-Konfidenzintervall (KI) Untergrenze Methode mit einem angenommenen wahren HR von 0,8 und einer Nicht-Unterlegenheits-Grenze von 1,08 (korrespondierend mit einer 60%-igen Retention des Effektes von Sorafenib gegen Plazebo), betrug die Power der Studie für die Aufdeckung der Nicht-Unterlegenheit etwa 97%. • Die Power der Studie um die Überlegenheit von Lenvatinib gegenüber Sorafenib zu zeigen, betrug etwa 82% unter Verwendung des Überlegenheits-Tests mit einem angenommenen wahren HR von 0,8. Basierend auf diesen Annahmen wurde die erforderliche Anzahl der Ereignisse (Todesfälle) im Per-Protocol-Analysis-Set (PPS; umfasst alle Patienten, die randomisiert wurden, mindestens eine Dosis der Studienmedikation erhalten und keine größeren Protokollabweichungen gezeigt haben) auf 666 geschätzt. Unter der Annahme, dass etwa 5% der Patienten mit schweren Protokollverletzungen und Patienten, die einer Behandlung zugewiesen wurden, aber keine erhalten hatten, vom PPS ausgeschlossen werden, sind im Full Analysis Set (FAS; umfasst alle Patienten, die randomisiert wurden) etwa

Item[a]	Charakteristikum	Studieninformation
		700 Ereignisse (Todesfälle) zum Zeitpunkt der Primäranalyse notwendig. Für die Schätzung wurden zwei Zwischenanalysen zur Futility berücksichtigt (bei circa 30% und 70% der Zielanzahl der Ereignisse). Es wurde geschätzt, dass etwa 940 Patienten (470 Patienten pro Behandlungsarm) randomisiert werden müssen, um 700 Ereignisse zu beobachten.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es gab zwei Zwischenanalysen: • Die erste Zwischenanalyse wurde am 13. Mai 2015 durchgeführt, nachdem etwa 30% der notwendigen Fallzahl (entsprach 249 Todesfällen) in beiden Behandlungsarmen beobachtet wurde. • Die zweite Zwischenanalyse wurde am 03. Februar 2016 durchgeführt, nachdem circa 70% der notwendigen Fallzahl (entsprach 521 Todesfällen) in beiden Behandlungsarmen beobachtet wurde. Basierend auf beiden Zwischenanalysen wurde empfohlen, die Studie ohne weitere Modifikationen fortzuführen.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung wurde mithilfe eines interaktiven Voice and Web Response Systems (IxRS) durchgeführt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden nach einem Zuteilungsverhältnis von 1:1 einem der zwei Behandlungsarme zugeteilt (Lenvatinib und Sorafenib). Die Stratifizierung erfolgte nach folgenden Kriterien (IxRS Stratifizierungsfaktoren): • Region: Region 1;(Asien-Pazifik), Region 2 (westliche Regionen: EU, Nordamerika, andere) • Makroskopische Invasion der Pfortader (MPVI) oder Extrahepatische Ausbreitung, oder beides: ja; nein • ECOG-PS: PS=0; PS=1 • Körpergewicht: < 60 kg; ≥ 60 kg
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Zuteilungssequenz erfolgte durch ein zentrales unabhängiges interaktives Voice / Web Response System (IxRS). Jedem potentiellen Studienteilnehmer wurde eine Screening-Nummer zugewiesen und mithilfe des IxRS Patienteninformationen registriert.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Zuteilung erfolgte zentral gemäß der durch das IxRS übermittelten Behandlung.
11	Verblindung	

Item[a]	Charakteristikum	Studieninformation
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Da es sich um eine offene Studie handelt, waren weder das Studienpersonal, noch die Patienten verblindet hinsichtlich der Studienbehandlung.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht relevant.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Alle Wirksamkeitsanalysen wurden auf Grundlage des FAS (primäre Analysen) bzw. des PPS (sekundäre Analysen) durchgeführt. **OS (Gesamtüberleben):** • Für das OS wurde ein Nicht-Unterlegenheits-Test von Lenvatinib gegenüber Sorafenib anhand eines zweiseitigen 95%-KIs bezüglich HR durchgeführt, mit einer Nicht-Unterlegenheits-Grenze von 1,08. Nicht-Unterlegenheit galt als erfüllt, falls bei der finalen Analyse die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag. Das HR und das entsprechende zweiseitige 95%-KI wurden mithilfe eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit der „Behandlung“ als Faktor, berechnet. • Für das OS wurde ein Überlegenheits-Test von Lenvatinib gegenüber Sorafenib anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, durchgeführt. Das Signifikanzniveau wurde dabei auf α=0,05 gesetzt. • Es wurden zwei Zwischenanalysen durchgeführt. Da die Studie nicht aufgrund einer der Zwischenanalysen gestoppt wurde, wurde die Nicht-Unterlegenheit bezüglich OS erstmals bei der finalen Analyse getestet, mit einer Nicht-Unterlegenheits-Grenze von 1,08. Dies zeigt, dass Lenvatinib mindestens 60% (korrespondierend zu δ=0,60) des Effekts von Sorafenib vs. Placebo – wie in der SHARP und Asia-Pacific-Studie beobachtet – aufgezeigt hat. Nicht-Unterlegenheit galt als erfüllt, falls bei der finalen Analyse die obere Grenze des zweiseitigen 95%-KI des HR nicht über dem Schwellenwert 1,08 lag. Falls Nicht-Unterlegenheit für den Endpunkt OS gezeigt wurde, sollte auf Überlegenheit getestet werden (korrespondierend zu δ=1). Überlegenheit galt als erfüllt, wenn mithilfe eines stratifizierten Log-Rank-Tests bei der finalen Analyse der zweiseitige p-Wert < 0,05 betrug. • Für das OS wurden der Median, das erste und dritte Quartil (Q1 und Q3) und die kumulative Wahrscheinlichkeit von OS zu ausgewählten Zeitpunkten mittels Kaplan-Meier (KM)-Schätzer

Item[a]	Charakteristikum	Studieninformation
		(Kaplan-Meier-Schätzer) für jeden Behandlungsarm berechnet und mit zweiseitigen 95%-KIs dargestellt. Die KM-Schätzer wurden für jeden Behandlungsarm über die Zeit geplottet. **PFS (Progression Free Survival):** • Für das PFS wurde die Differenz zwischen Lenvatinib und Sorafenib anhand eines stratifizierten Log-Rank-Tests, stratifiziert nach den IxRS Stratifizierungsfaktoren, bewertet. Das Signifikanzniveau wurde dabei auf α=0,05 gesetzt. • Für das PFS wurde das HR und das entsprechende zweiseitige 95%-KI mithilfe eines Cox-Proportional-Hazard-Modells, stratifiziert nach den IxRS Stratifizierungsfaktoren und mit der „Behandlung“ als Faktor, berechnet. • Für das PFS wurden der Median, Q1 und Q3 und die kumulative Wahrscheinlichkeit in 6-Monats-Intervallen mittels KM-Schätzer (Kaplan-Meier-Schätzer) für jeden Behandlungsarm berechnet und mit zweiseitigen 95%-KIs dargestellt. Die KM-Schätzer wurden für jeden Behandlungsarm über die Zeit geplottet. **TTP (Time to Progression):** • Für die TTP wurde die Differenz zwischen Lenvatinib und Sorafenib analog zum Vorgehen beim PFS bewertet. **ORR (Objective Response Rate):** • Für die ORR wurde die Differenz zwischen Lenvatinib und Sorafenib anhand eines zweiseitigen, stratifizierten Cochran-Mantel-Haenszel Chi-Quadrat-Test, stratifiziert nach den IxRS Stratifizierungsfaktoren, bewertet. Das Signifikanzniveau wurde dabei auf α=0,05 gesetzt. • Für die ORR wurde ein 95%-KI für jeden Behandlungsarm berechnet. • Für die ORR wurden die Odds Ratio und die Differenz in ORR mit einem 95%-KI für jeden Behandlungsarm dargestellt. **HRQoL:** • Deskriptive Analysen wurden basierend auf den patientenberichteten Outcomes (engl. patient-reported outcomes, PRO) der Fragebögen EORTC QLQ-C30, EORTC QLQ-HCC18 und EQ-5D, die während der Studienlaufzeit erhoben worden waren, durchgeführt. • Die Differenz zwischen den beiden Behandlungsarmen wurde im Längsschnitt und im Querschnitt untersucht. • Die sekundären Endpunkte des HCC-spezifischen Fragebogens EORTC QLQ-HCC18 und des EORTC QLQ-C30 wurden hierarchisch untersucht.

Item[a]	Charakteristikum	Studieninformation
		Sicherheit und Verträglichkeit: • Darstellung mittels deskriptiver Statistik auf Grundlage aller randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben (engl. Safety Analysis Sets, SAS). **Explorative Zielgrößen:** • Es wurden keine mehrfachen Anpassungen für die explorativen Zielgrößen verwendet. • Für DCR und CBR wurde die Differenz zwischen Lenvatinib und Sorafenib analog zum Vorgehen bei der ORR bewertet. • Für EORTC QLQC30, EORTC QLQ-HCC18 und EQ-5D wurden explorative Analysen durchgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	**Subgruppenanalyse:** • Es wurden Subgruppenanalysen für den primären Endpunkt OS sowie für den sekundären Endpunkt PFS durchgeführt • Die HR und die Mediane mit den dazugehörigen zweiseitigen 95 %-KI wurden für die Subgruppen, definiert durch die Stratifikationsfaktoren gemäß den Einträgen im IxRS sowie anderen Faktoren wie Alter, Geschlecht und Ätiologie (HBV, HCV und Alkohol und Alpha-Fetoprotein (AFP)) zu Baseline, dargestellt. **Sensitivitätsanalyse:** • Als Sensitivitätsanalyse bezüglich OS wurde mit einem nicht-stratifizierten Log-Rank-Test für den Vergleich der Behandlungsarme durchgeführt. Für die Schätzung des HRs wurde ein nicht-stratifiziertes Cox-Proportional-Hazard-Modell verwendet mit der Behandlung als einzigem Faktor. • Für OS wurde eine weitere Sensitivitätsanalyse durchgeführt nach Stratifizierungsfaktoren gemäß CRF. • Als Sensitivitätsanalyse bezüglich PFS wurde eine stratifizierte Analyse unter Verwendung aller Patienten mit einer Krankheitsprogressionen als Ereignis und allen verstorbenen Patienten durchgeführt. Dabei wurde ein Patient mit einer progressive disease (PD), auch wenn er fehlende Auswertungen hatte, eine neue Anti-Krebs-Therapie begonnen oder die Therapie aus anderen Gründen als eine PD abgebrochen hatte, oder ein verstorbener Patient nicht zensiert. • Als Sensitivitätsanalyse bezüglich TTP wurde eine stratifizierte Analyse unter Verwendung aller Patienten mit einer Krankheitsprogressionen als Ereignis durchgeführt. Dabei wurde ein Patient mit einer PD nicht zensiert, auch wenn er fehlende Auswertungen hatte, eine neue Anti-Krebs-Therapie begonnen oder die Therapie aus anderen Gründen als eine PD abgebrochen hatte.

Item[a]	Charakteristikum	Studieninformation
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) Anzahl aller randomisierten Patienten: N=954 • Lenvatinib: N=478 • Sorafenib: N=476 b) Anzahl der Patienten, die tatsächlich die geplante Behandlung bekommen haben: N=951 • Lenvatinib: N=476 • Sorafenib: N=475 c) Anzahl der Patienten, die in der Analyse des primären Zielkriteriums OS berücksichtigt wurden (FAS): N=954 • Lenvatinib: N=478 • Sorafenib: N=476
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Verlorene und ausgeschlossene Patienten nach Behandlungsarm: Lenvatinib, N=451: • Radiologische Progression: N=311 (65,1%) • Unerwünschtes Ereignis: N=63 (13,2%) • Klinische Progression: N=32 (6,7%) • Entscheidung des Patienten: N=28 (5,9%) • Lost-to-follow-up: N=3 (0,6%) • Widerruf der Zustimmung: N=9 (1,9%) • Andere: N=5 (1,0%) Sorafenib, N=451: • Radiologische Progression: N=347 (72,9%) • Unerwünschtes Ereignis: N=43 (9,0%) • Klinische Progression: N=33 (6,9%) • Entscheidung des Patienten: N=15 (3,2%) • Lost-to-follow-up: N=1 (0,2%) • Widerruf der Zustimmung: N=5 (1,1%) • Andere: N=7 (1,5%)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Am 01. März 2013 gab der erste Patient seine Einwilligung zur Studienteilnahme an der Phase-III-Studie. Die Patienten wurden dann laufend in die Studie aufgenommen. Der Datenschnitt für die Primäranalysen fand am 13. November 2016 statt. Der Zeitraum der Nachbeobachtung begann unmittelbar nach dem OFF-Treatment-Besuch, der innerhalb von 30 Tagen nach der Gabe der letzten Dosis erfolgte. Die Nachbeobachtung dauerte solange der Patient lebte, außer er widerrief seine Zustimmung, wobei die Nachbeobachtung

Item[a]	Charakteristikum	Studieninformation
		zum Gesamtüberleben alle 12 Wochen bei allen Patienten durchgeführt wurde, alle weiteren Krebsbehandlungen wurden notiert. Wenn ein Patient die Studie vor Woche 8 beendete, wurde der Scan für die Tumor-Bewertung sobald wie möglich nach Woche 8 durchgeführt, aber vor dem Beginn einer weiteren Krebsbehandlung. Patienten wurden vom Prüfarzt gemäß dem vorherrschenden, lokalen Therapiestandard behandelt. Patienten, die die Studie ohne Krankheitsprogression beendeten, erhielten weiterhin eine Tumor-Bewertung gemäß dem lokalen Therapiestandard.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.
a: nach CONSORT 2010		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

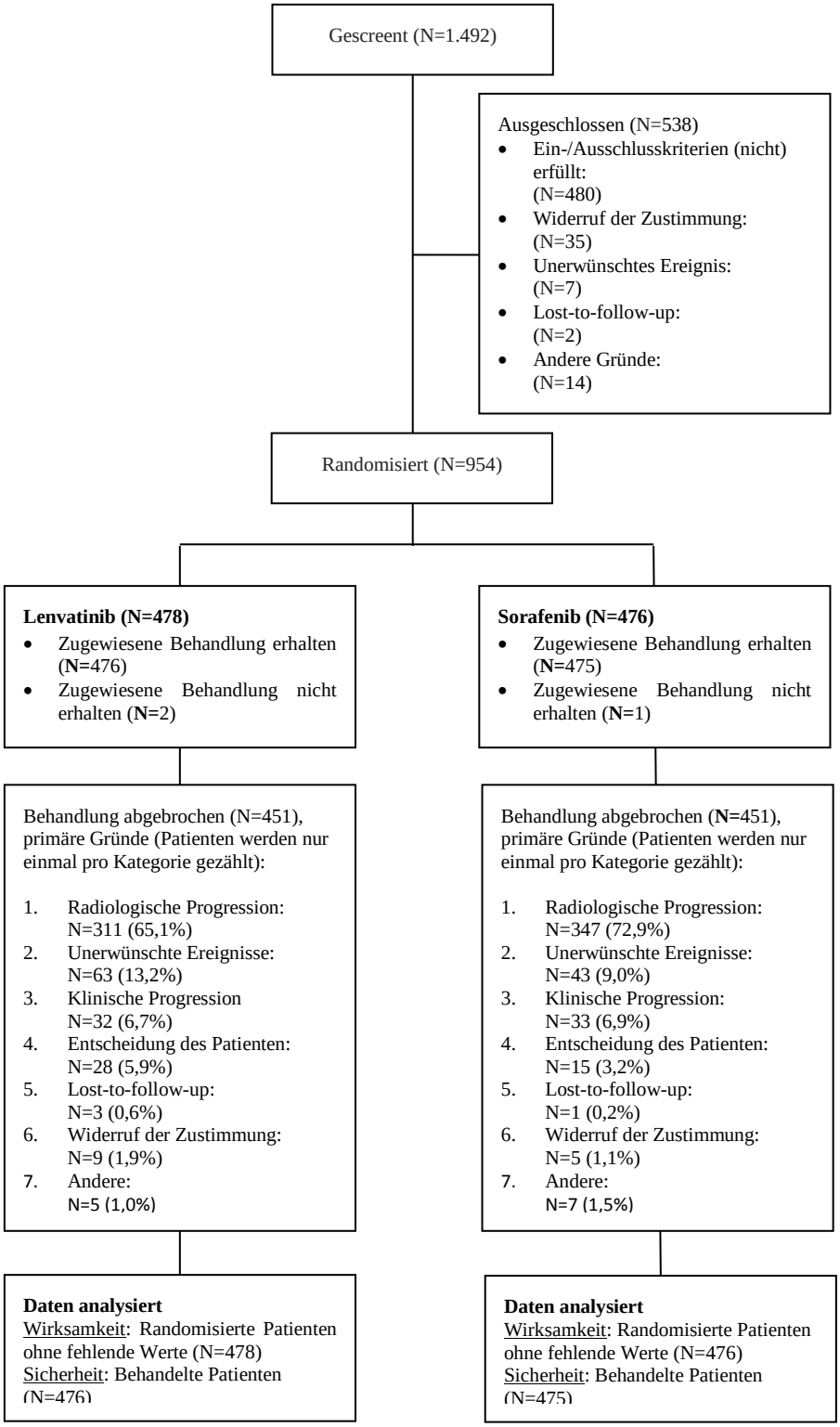


Abbildung 4-145: Flow Chart Studie 304

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.

- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C …) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Tabelle 4-85 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie 304

Studie: Studie 304

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Data on File: Clinical Study Report, E7080-G000-304	Studie 304

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Studie wurde randomisiert durchgeführt.

1.
<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

Die Zuteilung erfolgte rein zufällig mittels eines zentralen unabhängigen interaktiven Voice and Web Response Systems (IxRS).

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

2.
<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gruppenzuteilung erfolge durch ein zentrales unabhängiges interaktives Voice and Web Response System (IxRS).

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung.

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

In den Studienunterlagen finden sich keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Randomisierung wurde adäquat durchgeführt und die Gruppenzuteilung erfolgte rein zufällig mittels eines zentralen unabhängigen interaktiven Voice and Web Response Systems (IxRS). Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren oder - wie bei den Endpunkten zu Sicherheit und Verträglichkeit und den Subgruppenanalysen - gemäß Verfahrensordnung des G-BA berichtet wurden (z.B. die Berechnung der Effektmaße für den Vergleich der Behandlungsgruppen). Die Darstellung der Ergebnisse gibt keinen Anlass zur Vermutung, dass Ergebnisse selektiv berichtet wurden. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschließt.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Gesamtüberleben (Overall Survival, OS)

1. **Verblindung der Endpunkterheber**

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Für die Auswertung wurde das „Full Analysis Set“ (auch als „Intention-To-Treat [ITT] Analysis Set“ oder ITT-Population bezeichnet) berücksichtigt. Dies schließt alle Patienten ein, die randomisiert wurden. Somit wurde das ITT-Prinzip eingehalten.

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts OS.

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotential wird als niedrig eingestuft, da die Beeinflussung hinsichtlich des Ergebnisses des Endpunkts OS auch bei einem Open-Label-Studiendesign durch das Wissen um die Behandlung weder durch den Patient, noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im statistischen Analyseplan SAP definiert waren. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts OS als niedrig eingestuft

Endpunkt: Progressionsfreies Überleben (Progression Free Survival, PFS, IIR Bewertung)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie. Die Bewertung des Endpunkts erfolgte durch ein verblindetes unabhängiges Reviewboard (IIR).

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Für die Auswertung wurde die ITT-Population berücksichtigt. Dies schließt alle Patienten ein, die randomisiert wurden. Somit wurde das ITT-Prinzip eingehalten

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts PFS.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotential wird als niedrig eingestuft, da der Endpunkt PFS mit klar definierten, objektiven Kriterien erfasst wurde (mRECIST und RECIST v.1.1). Wegen der Bewertung der Progression durch ein verblindetes unabhängiges Reviewboard ist keine Beeinflussung aufgrund des Open-Label-Studiendesigns zu erwarten. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolge auf Basis der ITT-Population, die alle randomisierten Patienten einschließt. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts PFS (IIR Bewertung) als niedrig eingestuft.

Morbidität und Gesundheitsbezogene Lebensqualität
Endpunkt: EORTC QLQ-C30

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei diesem subjektiv erhobenen Endpunkt ist der Endpunkterheber der Patient selbst. Da die Patienten in der Studie nicht verblindet waren, liegt keine verblindete Endpunkterhebung bezüglich des EORTC QLQ-C30 vor.

2. Adäquate Umsetzung des ITT-Prinzips

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Für die Auswertung wurden alle Patienten mit Werten zu Baseline und mindestens einem Folgewert der ITT-Population berücksichtigt. Der Anteil berücksichtigter Patienten lag knapp unterhalb von 95%,aber über 70%.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts EORTC QLQ-C30.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95%. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-C30 als hoch eingestuft.

Endpunkt: EORTC QLQ-HCC18

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei diesem subjektiv erhobenen Endpunkt ist der Endpunkterheber der Patient selbst. Da die Patienten in der Studie nicht verblindet waren, liegt keine verblindete Endpunkterhebung bezüglich des EORTC QLQ-HCC18 vor.

2. Adäquate Umsetzung des ITT-Prinzips

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Für die Auswertung wurden alle Patienten mit Werten zu Baseline und mindestens einem Folgewert der ITT-Population berücksichtigt. Der Anteil berücksichtigter Patienten lag knapp unterhalb von 95%, aber über 70%.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts EORTC QLQ-HCC18.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95%. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-C30 als hoch eingestuft.

Endpunkt: EQ-5D VAS

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Bei diesem subjektiv erhobenen Endpunkt ist der Endpunkterheber der Patient selbst. Da die Patienten in der Studie nicht verblindet waren, liegt keine verblindete Endpunkterhebung bezüglich der EQ-5D VAS vor.

2. Adäquate Umsetzung des ITT-Prinzips

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Für die Auswertung wurden alle Patienten mit Werten zu Baseline und mindestens einem Folgewert der ITT-Population berücksichtigt. Der Anteil berücksichtigter Patienten lag knapp unterhalb von 95%, aber über 70%.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts EQ-5D VAS.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Insgesamt wird das Verzerrungspotenzial als hoch eingestuft, da mehr als 70%, aber weniger als 95% der Patienten der ITT-Population in die Analysen aufgenommen wurden. Jedoch fielen die Rücklaufquoten hoch aus und lagen deutlich über 70% und nur knapp unter dem Zielwert von 95%. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden. Die Beeinflussung hinsichtlich des Ergebnisses durch das Wissen um die Behandlung kann nicht ausgeschlossen werden, erhöht aber das Verzerrungspotenzial nicht weiter. Es lagen keine weiteren Aspekte für ein hohes Verzerrungspotenzial vor. Somit wird das Verzerrungspotenzial für den Endpunkt EORTC QLQ-C30 als hoch eingestuft.

Sicherheit und Verträglichkeit
Endpunkt: Unerwünschte Ereignisse (UE)

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Endpunktauswertung erfolgte auf Grundlage des Sicherheitsanalysesets (auch als Safety Analysis Set (SAS) bezeichnet). Dieses schloss alle Patienten ein, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl aufgetreten. Der Nichtberücksichtigungsanteil in der Auswertung lag unter 5%. Somit wurde das ITT-Prinzip eingehalten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts UE.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotential der UE wird als niedrig eingestuft, da eine Beeinflussung hinsichtlich des Ergebnisses der Endpunkte auch bei offenem Studiendesign durch das Wissen um die Behandlung weder durch den Patient noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis des SAS, in das alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhalten haben. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor.

Endpunkt: Schwerwiegende UE (SUE)

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Endpunktauswertung erfolgte auf Grundlage des SAS. Dieses schloss alle Patienten ein, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl aufgetreten. Der Nichtberücksichtigungsanteil in der Auswertung lag unter 5%. Somit wurde das ITT-Prinzip eingehalten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts SUE.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotential der SUE wird als niedrig eingestuft, da eine Beeinflussung hinsichtlich des Ergebnisses der Endpunkte auch bei offenem Studiendesign durch das Wissen um die Behandlung weder durch den Patient noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis des SAS, in das alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhalten haben. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor.

Endpunkt: UE, die zum Therapieabbruch führten

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Endpunktauswertung erfolgte auf Grundlage des SAS. Dieses schloss alle Patienten ein, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl aufgetreten. Der Nichtberücksichtigungsanteil in der Auswertung lag unter 5%. Somit wurde das ITT-Prinzip eingehalten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts UE, die zum Therapieabbruch führten.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotential der UE, die zum Therapieabbruch führten, wird als niedrig eingestuft, da eine Beeinflussung hinsichtlich des Ergebnisses der Endpunkte auch bei offenem Studiendesign durch das Wissen um die Behandlung weder durch den Patient noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis des SAS, in das alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhalten haben. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor.

Endpunkt: UE mit CTCAE Grad ≥ 3

1. Verblindung der Endpunkterheber

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelt sich um eine Open-Label Studie.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Endpunktauswertung erfolgte auf Grundlage des SAS. Dieses schloss alle Patienten ein, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl aufgetreten. Der Nichtberücksichtigungsanteil in der Auswertung lag unter 5%. Somit wurde das ITT-Prinzip eingehalten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise für eine ergebnisgesteuerte Berichterstattung des Endpunkts UE mit CTCAE Grad ≥ 3

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise für sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotential der UE mit CTCAE Grad ≥ 3wird als niedrig eingestuft, da eine Beeinflussung hinsichtlich des Ergebnisses der Endpunkte auch bei offenem Studiendesign durch das Wissen um die Behandlung weder durch den Patient noch durch den Arzt zu erwarten ist. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da alle Analysen prospektiv im SAP definiert waren. Die Analyse erfolgte auf Basis des SAS, in das alle Patienten aufgenommen wurden, die mindestens eine Dosis der Studienmedikation erhalten haben. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor.